Development of a Sound Projection Prototype

Master Thesis

Author: Ben-Daniel Keller

April 2015

University of Music and Performing Arts Graz
Institute of Electronic Music and Acoustics

Graz University Of Technology

Advisors:

DI Dr. rer. nat. Franz Zotter

DI Hannes Pomberger

Assessor:

o.Univ.-Prof. Mag. art. DI Dr. techn. Robert Höldrich

Abstract:

This master thesis describes the development and construction of a novel loudspeaker array that creates highly directional sound radiation on a spherical section. This section is flanked by sound hard boundaries which confine radiation to the spherical section. Limiting possible radiation to a spherical section strongly increases the spatial resolution, while the number of loudspeakers is the same as with a full-sphere array. The beam pattern is composed of harmonic functions especially derived for this spherical section. A suitable combination of these harmonics results in beamforming. The ambition is to aim these sound beams against the wall to excite directed reflection paths. In analogy to video projection the point of acoustic impact should be perceived as the origin of a spherical wave. For this purpose, wall reflection must preferably be diffuse. The finite-length boundary's influence on beam steering will be investigated. Is it possible to use this sonic projection instrument as a means of 2D spatialisation if the wall reflection is only diffuse?

Kurzfassung: Entwicklung des Schallprojektors

Im Rahmen dieser Masterarbeit wurde ein Lautsprecher-Array entwickelt und gebaut, das auf einem Kugelausschnitt stark bündelnd abstrahlt. Der Kugelausschnitt wird von schallharten Flächen flankiert, um nur auf diesem für eine geführte Schallausbreitung zu sorgen. Durch die Beschränkung der Abstrahlung auf diesen Kugelausschnitt gelingt es, bei gleichbleibender Lautsprecheranzahl die Auflösung im Vergleich zu einem Kugelarray enorm zu erhöhen. Das erreichbare Bündelungsmuster (beam pattern) setzt sich aus speziell für diesen Kugelausschnitt hergeleiteten Harmonischen zusammen. Zur Bündelungssteuerung (beamforming) werden die Harmonischen geeignet kombiniert. In der Anwendung wird die mit dem Array erzeugte Bündelung auf eine Wand ausgerichtet, um dort gezielte Reflexionswege anzuregen. In Anlehnung an die Videoprojektion – nur in diesem Fall akustisch – soll die getroffene Stelle wahrnehmbarer Ausgangspunkt einer Kugelwelle sein. Dazu muss die Wandreflexion möglichst diffus sein. Erlaubt eine nur beschränkt diffuse Wandreflexion immer noch den Einsatz des Schallprojektors als neues 2D-Klangverräumlichungsinstrument?

Acknowledgements

This work was mostly done in one year's time at the IEM's offices and laboratories. Therefore, I want to thank the whole team of the IEM for their support and the nice atmosphere at work. Most of all, I want to thank my advisors:
Thank you, Franz Zotter, for suggesting to me this interesting research topic, which gave me the chance to develop something really new. I want to thank you for the constant encouragement, help and explanation, and your catching enthusiasm for research.
Thank you, Hannes Pomberger, for the help with mathematics and programming, the prior research in this field, the help with the measurements, and for finding within minutes my ridiculous coding mistakes (which kept me busy for hours).

I further want to thank:
Matthias Frank, for advanced troubleshooting sessions w.r.t. PureData, Matlab, LaTex, and the institute's hardware;
Stefan Warum, for giving helpful tips concerning design and construction and for organizing all the hardware orders;
Robert Höldrich, for his lecture on musical acoustics that made me start my studies at the IEM in the first place.

I want to thank my girlfriend Liselotte for her patience with my graduation and the delay of long overdue holidays. I am grateful for her encouragement and her witty, humorous summaries of my project. Also, her diligent proof-reading once more showed me that English is easy to learn and difficult to master.

I am deeply grateful to my parents, who always supported my education.

I want to thank the Wirtschaftskammer Österreich for granting me the 'Wissenschaft trifft Forschung' Scholarship.

Last but not least, I want to thank the people of Austria for financing my studies.

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Graz, ...

(Signature)

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am..................... ....

(Unterschrift)

Contents

Notations

SH	spherical harmonics
SSH	spherical sector harmonics
ICO	IEM'S icosahedron loudspeaker array
ϑ	zenith angle
φ	azimuth angle
c	speed of sound in air (unless specified $c = 343.2$ m/s is assumed)
f	frequency
$k = \omega/c = 2\pi f/c$	wave number
λ	wavelength
a	array radius
r	projection radius
r_0	microphone radius
N	maximum order of modal decomposition
ν	order of harmonic
μ	degree of harmonic
j	unit imaginary number, such as $j^2 = -1$
$\Re\{.\}$	real part
$\mathbf{r} \equiv (r, \varphi, \vartheta)$	position vector in spherical coordinates
	with φ the azimuth angle in $[0, 2\pi]$ and ϑ the zenith angle in $[0, \pi]$.
	where ϑ is measured from the polar axis $\mathbf{z}$
$\boldsymbol{Y}_n^m(\varphi, \vartheta)$	spherical sector harmonic of order n and degree m evaluated at φ, ϑ
$h_n^{(2)}(x)$	spherical Hankel function of the second kind of order n
$p(\mathbf{r}, t)$	(instantaneous) sound pressure at position $\mathbf{r}$, in the time domain
$p(\mathbf{r}, k)$	(complex) sound pressure at position $\mathbf{r}$, in the wave number domain
$c_{nm}(k)$	SSH wave spectrum of the radiating field, independent of any radius r
$\zeta_{nm}(r, k)$	SSH spherical wave spectrum of surface particle velocity z

1 Introduction

Only sources emitting strongly directional sound beams can separately excite the distinct specular wall reflections of a room. If the reflection of a beam is diffuse enough, it will be the origin of a secondary spherical wave at the point of impact. This is similar to diffuse optical reflection utilized for the projection of images on cinema screens. Fig. 1 shows the functional principle.

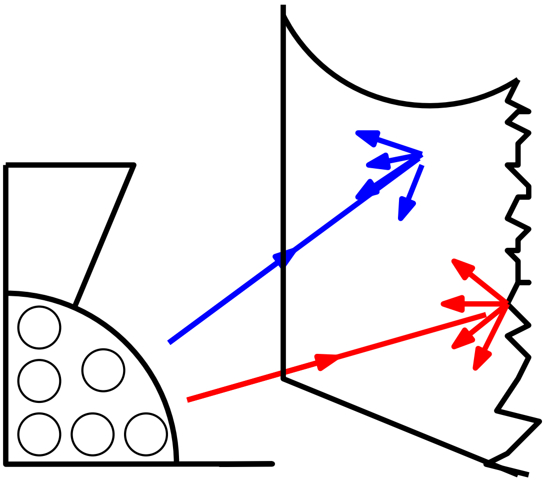
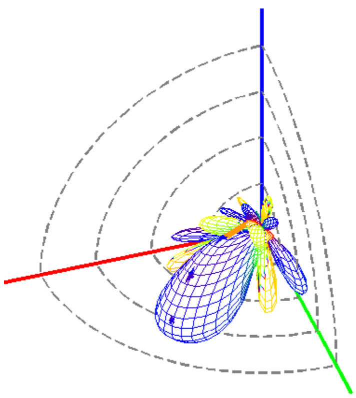


Figure 1: Principle of the sound projector exciting diffuse wall reflections

A musical utilization of specular acoustic reflections was impressively demonstrated in concerts using IEM's Icosahedron loudspeaker array (ICO) employing third-order spherical harmonic beamforming [SZF14, ZFFR14]. Neither has the ICO been exploited for exciting diffuse reflections nor is its directivity sufficient to produce diffuse sound louder than the direct sound reaching a large audience.
The interesting range of directions for sound projection is limited to the acoustic screen, i.e., the diffusely reflecting wall. According to recent work of Pausch and Pomberger [Pau13, PP14], sound hard bounding surfaces can be inserted to narrow down the angular beamforming range and to get an increased spatial definition within.
Pursuing the novel idea of sound projection to a diffusely reflecting acoustic screen, this thesis presents a prototype and the theoretical basis tailored to the goal. In particular, the design of the layout, the derivation of underlying mathematical functions and their optimal sampling are presented in this work. The fully equipped 20 channel prototype of the sound projector is depicted in fig. 2. Fig. 3 shows the measured results for beampattern near the

aliasing frequency. The main lobe width of the beampattern is comparable to one of an 11th order full-sphere array, which would require at least 144 loudspeakers.

Figure 2: Prototype of the sound projector equipped with 20 transducers



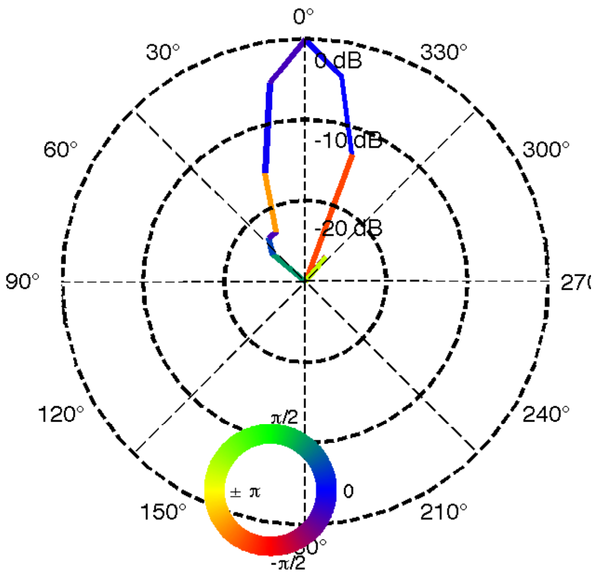


Figure 3: Measurement results for beamforming evaluated at $f = 3$ kHz and measured at $r = 0.8$ m

1.1 Overview

A short overview of the previous work in the research fields of loudspeaker arrays and guided radiation is given in chapter 2.
Chapter 3 shows variants for the array design, especially for its cabinet. Acoustic criteria, such as the maximization of the cabinet's volume, are considered . While keeping a compact size to support the mobility of the device the length for the bounding surfaces are maximized.
In chapter 4 modal solutions for the wave equation are derived for this special case: Spherical sector harmonics (SSH). Corresponding radial focussing filters are also explained in this chapter.
The positioning of the loudspeakers on the spherical surface is crucial to obtain a numerically reasonable SSH-based beampattern control (modal beamforming). Sampling strategies and optimization are explained in chapter 5.
After construction, acoustic measurements were made to obtain the array's directional impulse responses. The measurement results provide information about beamforming controllability. A description of the measurement and the evaluation is given in chapter 7.
The functionality is further examined with a real-time interface for beamforming. Various tests with the array are described in chapter 8.

2 Previous Work

2.1 Loudspeaker Arrays

The synthesis of various sound radiation patterns has been of interest for quite a long time. Every musical instrument has a distinct character of sound radiation. The acoustical reasoning behind this goal is that a loudspeaker array able to produce resembling patterns would no longer be perceived as a loudspeaker failing to reproduce the authentic characteristics of a live-played musical instrument. A successful implementation should make the rich perceivable spatial details synthetically reproducible in a concert hall. Moreover, there has always been an interest besides reproducing existing instrumental sound radiation patterns. Highly directional sound beams are used artistically in electronic music [SZF14]. Also, adjustable directivity is used in room acoustic measurements to achieve directional room impulse responses [Beh06, KPPV14].

Spherical loudspeaker arrays allow for full spherical directional control. A dodecahedron loudspeaker array was introduced at IRCAM [WDC97]. Further investigations on various spherical loudspeaker arrays were done at CNMAT [KW04] and an array featuring 120 loudspeakers was introduced by [AFKW06].

The IEM also has longtime experience with multichannel holophonic sound reproduction systems. A icosahedron loudspeaker array was introduced by [ZSH07]. Because the shared volume for the loudspeakers causes crosstalk, Lösler introduced novel parallel IIR filters for efficient MIMO crosstalk cancellation in his master thesis [Lös14].

Kronlachner developed cross-platform Ambisonics processors in his master thesis [Kro14] (available in the ambix plug-in suite), which make the ICO a versatile (musical) spatialisation instrument. It is able to stimulate rich acoustic room reflections like a real instrument and differently from standard loudspeakers. However, the ICO has some limitations:

- Limited directivity: the maximum order of the produced beams is $N = 3$
- Full directivity is only achieved in a limited frequency range:
  - *Lower frequency boundary*: the limiting factors are the array size, the loudspeaker's maximum linear excursion and its maximum power

– *Upper frequency boundary for spatial aliasing*: depends on order N and radius r according to $N > k\,r$

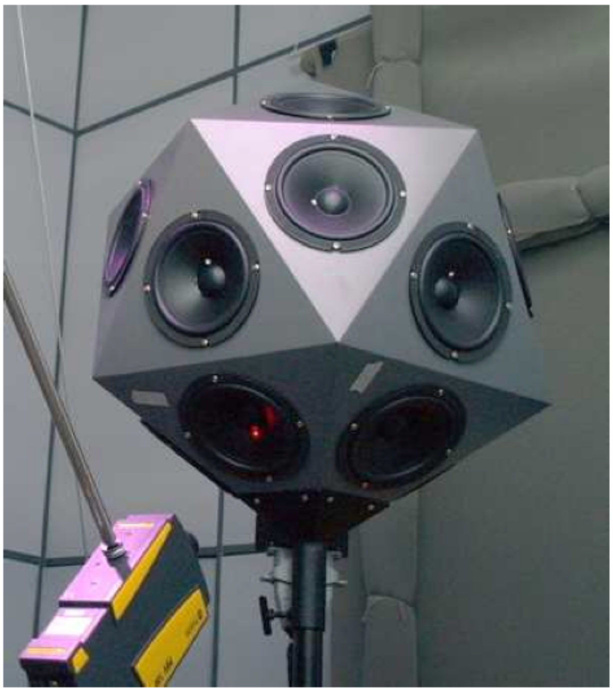


Figure 4: IEM's icosahedral loudspeaker array (ICO) [ZH07]

The IEM's experience with the ICO's ability to excite distinct wall reflections created the idea of a loudspeaker array with an even higher directivity on a restricted angular range. This would allow for further investigations and possible sound projection to diffuse reflecting walls. The novel prototype is designed exactly for that purpose and further allows to examine Pomberger's theory of SSH beamforming within angular boundaries [PZ13].

2.2 Sound Field Within Angular Bounding Surfaces

Modal beamforming with conventional spherical arrays is based on the decomposition of sound pressure into spherical harmonics (SH) [ME02]. This requires a distribution of loudspeakers covering all directions - even if not all directions of steering are of interest. In cases where not all the steering range is needed, a reasonable spatial resolution comes at a large cost. On a full sphere and critical sampling the maximum achievable spatial resolution calculates as $N \leq \sqrt{L} - 1$, where N is the maximum order of SH achievable and L is the number of loudspeakers. The ICO with its 20 loudspeakers only achieves directivity of order $N = 3$. In geodesy SH are extensively used, wherefore their enhancement, the spherical cap harmonics (SCH), were first introduced there [HC97]. SCH are comparable to SH, with the difference that the zenith solution calculates using non-integer orders ν in the associated Legendre functions. Taking up this idea, Pomberger introduced a more efficient sound field decomposition into SCH especially derived for the geometry [Pom10, PZ13]. Focussing on the area of interest increases spatial efficiency: Either the number of sensors can be limited maintaining the directivity or directivity is increased using the same number of sensors. However, it must be guaranteed that no sound arrives from directions outside this area. This can only be achieved theoretically employing infinitely long, sound-hard bounding surfaces for the sound field. In his master thesis at IEM Kößler investigated the insertion of a rigid cone as a diffraction object for the microphone case and derived the SCH for the geometry [Köß11].

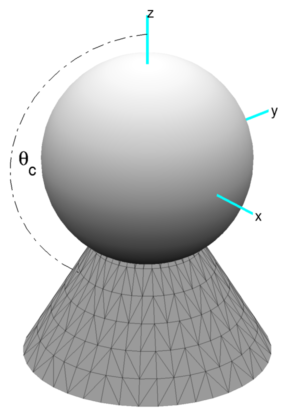


Figure 5: Geometry of the IEM's spherical cap microphone array introduced by [Köß11]

Pomberger and Pausch (Master alumnus at IEM) [Pau13, PP14] introduced a rigid double cone microphone prototype. Because it was intended to only record panoramic acoustic scenarios, there are two boundaries in the zenith angle yielding a double cone. Any pressure distribution on the surface area can be decomposed into Spherical Segment Harmonics. These harmonics are calculated using non-integer orders ν for the associated Legendre functions of the zenith solution. The array holds 64 microphone capsules that allow decomposition into 64 harmonics up to order $N = 9.9$.

Figure 6: The IEM's panoramic microphone array prototype built by Pausch [Pau13]

3 Design of the Louspeaker Cabinet

3.1 Prior Considerations and Important Aspects

The planned array consists of a spherical sector holding the loudspeakers and four bounding surfaces at a left and right azimuth angle and a top and bottom zenith angle. These surfaces end at a finite radius. To design the limiting angles, the radius of the array, and the extent of the angle limiting surfaces the following main aspects have to be considered:

- **Angular projection range:**
  The target application employs the sound projection on the screen of a video projection. Therefore the azimuth and zenith angular aperture of the bounding surfaces (φ, ϑ) were chosen to match the classic screening ratio for video projection of $4:3$. The array is best positioned in front of the audience, facing the wall. A reasonable distance to the reflecting wall demands an azimuthal aperture of $\varphi = 90°$ which leads to a zenithal aperture of $\vartheta = 67.5°$.

- **Length of bounding surfaces:**
  Ideally the bounding surfaces should be of infinite length, which would prohibit radiation of sound to the backside of the array. Obstacles bigger than half of the wavelength λ cause shadowing rather than diffraction ($l_b = \lambda/2$). Thus, the length of the bounding elements corresponds to the lower frequency boundary for directed sound radiation. The targeted prototype will not be a permanent installation. To minimize space requirements for storage or transport the length was chosen with $l_b = 550\,\text{mm}$, yielding a *lower frequency limit* $f_{l,b} = 320\,\text{Hz}$.

- **Array radius, achievable resolution and number of loudspeakers:**
  Velocity on the spherical surface is controlled at discrete locations of the loudspeakers. Similarly to the aliasing frequency with discrete time sampling, the number and the locations of the loudspeakers on the sphere lead to a spatial sampling criterion. Following [Wil99], the upper *frequency limit for spatial aliasing* $f_{u,alias}$ decreases with increasing array radius a (overall depending on the maximum order of modal decomposition N) according to the relation

$$N > ka, \tag{3.1}$$

$$f_{u,alias} < \frac{Nc}{2\pi a}, \tag{3.2}$$

with $k = \frac{\omega}{c} = \frac{2\pi f}{c}$ denotes the wave number. W.r.t. spatial aliasing a small radius a with a high number of loudspeakers is required. On the other hand, directed radiation and a small array radius requires large membrane excursions at low frequencies. The maximum linear excursion of a loudspeaker then becomes the limiting factor. Increasing the array radius brings relief here.

3.2 Choice of Loudspeakers

Due to their performance, a set of relatively small yet powerful loudspeakers (membrane diameter $= 37\,\mathrm{mm}$) with a convincing maximum linear excursion ($x_{max} = \pm 4\,\mathrm{mm}$) were chosen. The decision to use the same amplifier that is recently used with the ICO (sonible d24) limited the maximum number of channels to 24. However, the SSH derived for the geometry (see the following chapter 4) allow for an easy order truncation at a maximum order $N = 10.6$ supplying 20 harmonics (see fig. 13). The choice was to use 20 loudspeakers to control a maximum of equally many harmonics as the array carries loudspeakers (critical sampling). After optimal positioning of the loudspeakers (see chapter 5) the array radius was chosen to be as small as possible with $a = 235\,\mathrm{mm}$.

Figure 7: Chosen loudspeaker

3.3 Expected Performance

With the decisions made for the array radius $a = 235\,\mathrm{mm}$ and the resulting maximum order $N = 10.6$ a few estimations concerning the array's performance can already be made.

- with a, eq. (3.2) provides an upper frequency limit for spatial aliasing of $f_{u,alias} = 2.47$ kHz.

- With a and the orders calculated in chapter 4.1.2, the required radial filters (see chapter 4.2) can be calculated. High orders normally have to be amplified strongly at low frequencies to attain directivity at a larger distance. Below a certain frequency the required amplification for a certain order ν exceeds a 30 dB range. (The gain limit is usually introduced due to positioning precision and matching errors. Without any gain limit such small errors would be amplified enormously.) The gain limit yields lower frequency limits for each order ν. For example, the harmonic of order $\nu = 1.97$ can only be used fully for frequencies $f_{low,1.97} > 150Hz$. Below this frequency the order dependent gain has to be limited.

The expected frequency boundaries w.r.t array geometry are depicted in fig. 8.

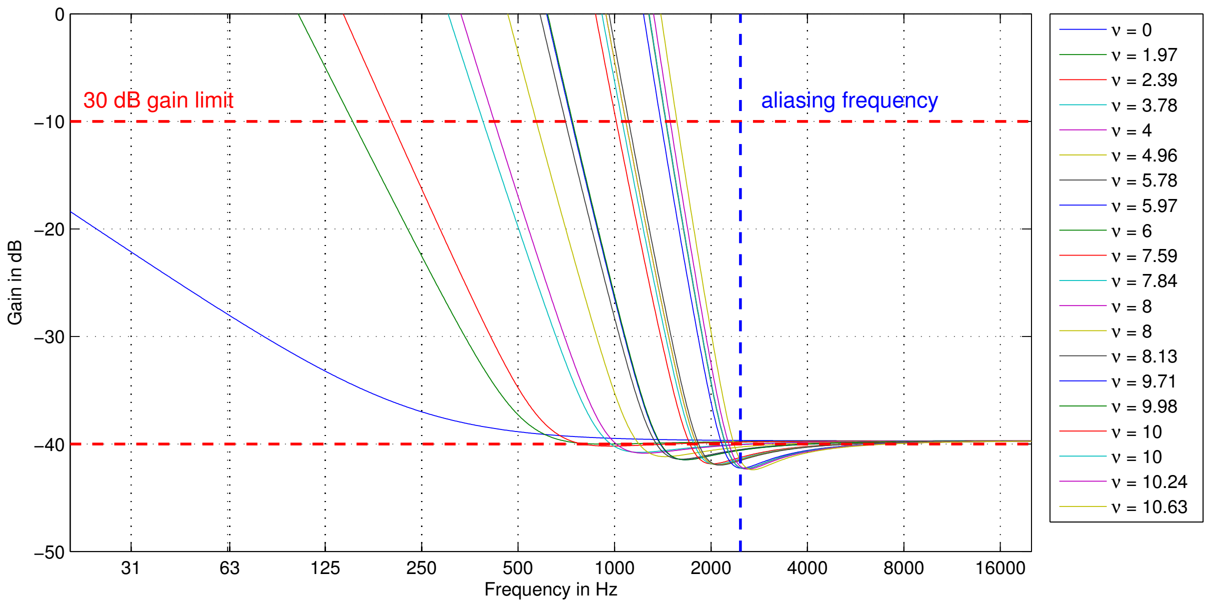


Figure 8: Radial filters and frequency limits depending on the array's geometry

Note:

The loudspeakers share a common volume. When loudspeakers are operated in-phase (which is the case with order $\nu = 0$), the cabinet shows a higher acoustical stiffness than in out-of-phase operation (which is the case with orders $\nu > 0$). At other operation modes than in-phase, the common enclosure will decrease the required currents, transduced to force on the loudspeaker membranes, to reach a desired excursion. This means there is no constant ratio of electrical current to excursion $(\frac{i(t)}{x(t)})$ for all operation modes.

3.4 Design Variants

There are numerous possible designs for the loudspeaker cabinet that contain the spherical sector and four angular bounding surfaces. The design process was done in the AutoCAD Inventor 3D CAD software. It allows to develop and adjust geometries in every step of the design process.

- **Initial design idea: vertical symmetry, big cabinet**
  The initial idea was a shoebox-shaped cabinet with the four funnel-like angular bounding surfaces lancing in one lateral face, ending with the spherical sector surface. The cabinet is vertically symmetric. Possible steering directions are up- and downward. The plane for vertical symmetry equals the equator of the spherical surface. The funnel-like bounding surfaces and the spherical sector are designed similarly to the model depicted in fig. 9. However, they are boxed in a cuboid cabinet with a large volume.

- **First design modification: vertical symmetry, small cabinet**
  To facilitate transportation by smaller outer dimensions the idea was to design the cabinet in a different way: The center of the sphere is positioned on one edge of the box. The azimuthal boundaries then equal two of the box's lateral faces. The equator of the spherical surface equals the plane of vertical symmetry. Possible steering directions are up- and downward. The volumes inside the top and bottom conical surfaces as well as the spherical surface are coupled, thus providing sufficient acoustic volume for loudspeaker operation. The design of the first modification is illustrated in fig. 9.

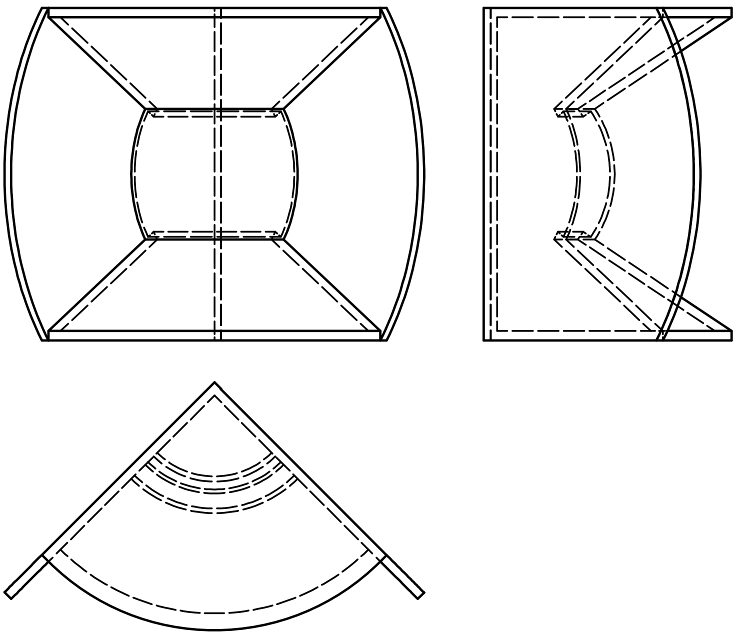


Figure 9: Possible loudspeaker cabinet variant: vertical symmetry, small cabinet

- **Second design modification: upward steering only, small cabinet**
  Positioning the sound projector on the floor, only upward beam steering is needed. (Mounting the sound projector on the ceiling, the beam steering is aimed downward.) The main difference from prior design considerations is that the lower half of cabinet and spherical sector is cut off. A flat bounding surface is inserted instead of the prior plane of vertical symmetry. A sketch is depicted in figure 10. When the sound projector is positioned directly on the floor, it necessarily has to be in front of the audience. Further, the floor acts as an extension to the bottom bounding surface. Setting up the array in the corner of a room increases the effective length of three bounding surfaces, which might be beneficial w.r.t. beamforming capabilities. However, the reduced complexity of the bounding surfaces is the main design benefit: The bottom and azimuth bounding surfaces are planes, only the upper azimuth boundary is a cone.

The array's overall size is even further reduced for transportation: the bottom wall is detachable and the side walls are mounted on hinges and can easily be folded inward. This also serves as protection for the loudspeakers during transportation. The CNC milling machine, which is used for the manufacturing of the prototype, works on imported CAD

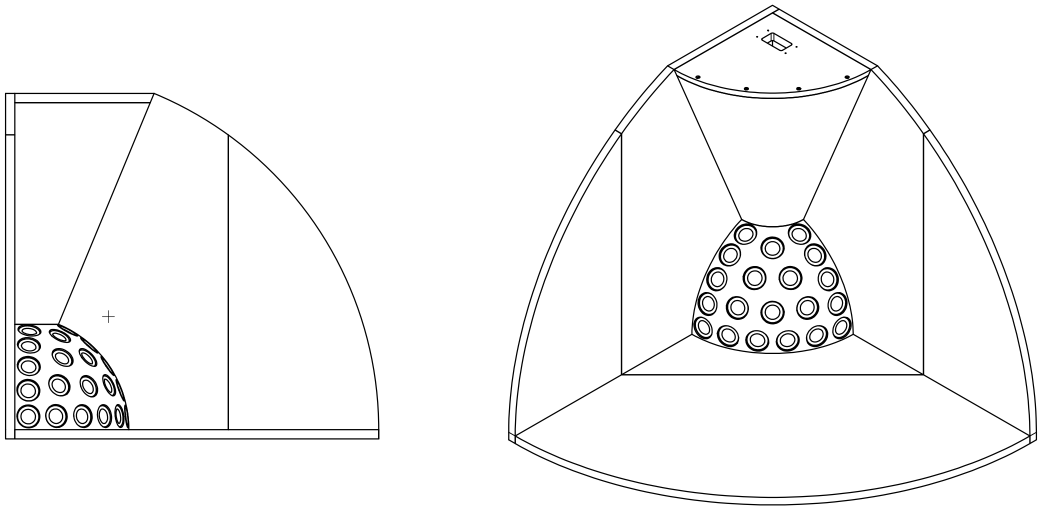


Figure 10: Sketch of the chosen cabinet design: the spherical equator coincides with the bottom boundary

data. Therefore every single drilling, as well as the cavity for the connection port in the top board and for the cabinet flange in the bottom board, is included in the CAD model. The remaining processing steps reduce to finish and assembly. Fig. 11 shows the CAD rendered design.

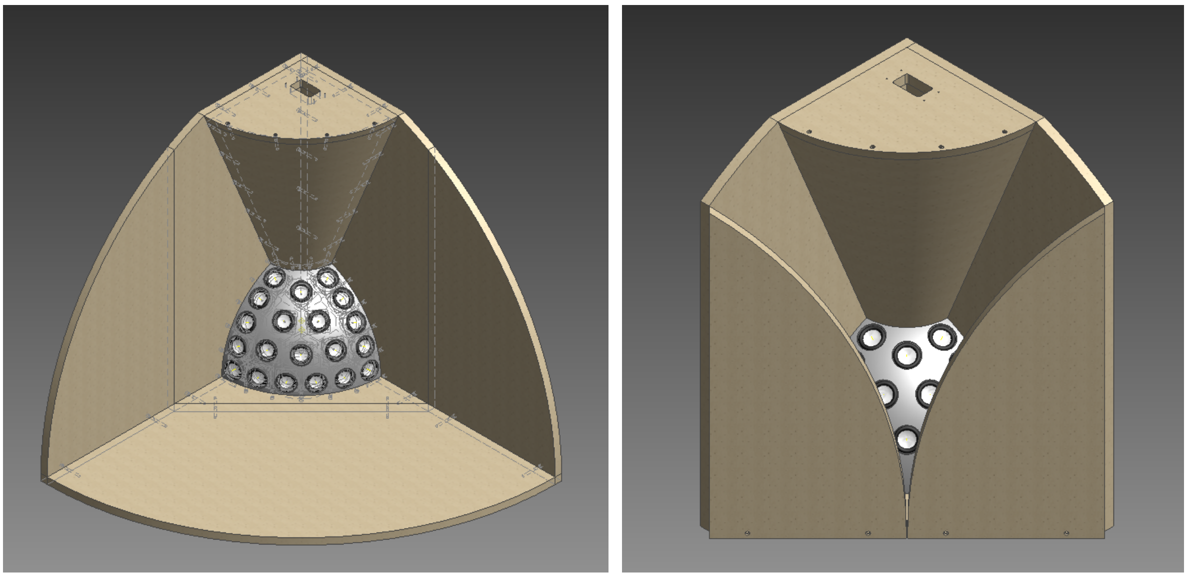


Figure 11: CAD rendered pictures of the prototype:
(left) fully assembled with all drillings visible, (right) partly disassembled for transportation

4 Solution of the Helmholtz Equation

The beam pattern of radiation is composed of harmonic functions especially derived for the acoustics of this spherical section. The Helmholtz equation describes the sound pressure variation $p(\mathbf{r}, \omega)$ in a homogeneous acoustic field with no viscosity. Therefore it needs to be solved for the geometry. The Helmholtz equation writes as [Wil99, Zot09]:

$$(\Delta + k^2)\, p(\mathbf{r}, \omega) = 0, \tag{4.1}$$

where Δ denotes the Laplace operator and $k = \omega/c$ the wave number that contains the frequency variable $\omega = 2\pi f$. The position vector $\mathbf{r} = (r, \varphi, \vartheta)^T$ is defined in the spherical coordinate system. (For a detailed definition of the spherical coordinate system used here see appendix A.) The dependency on the frequency variable ω is omitted in the following derivations. The Laplace operator for a function f in spherical coordinates writes as

$$\begin{aligned} \Delta f &= \Delta_r f + \Delta_\vartheta f + \Delta_\varphi f \\ &= \frac{1}{r^2}\frac{\partial^2 r^2 f}{\partial r^2} + \frac{1}{r^2 \sin^2(\vartheta)}\frac{\partial^2 f}{\partial \varphi^2} + \frac{1}{r^2 \sin\vartheta}\frac{\partial}{\partial \vartheta}\left(\sin\vartheta \frac{\partial f}{\partial \vartheta}\right). \end{aligned} \tag{4.2}$$

Combining eq. (4.2) and eq. (4.1), the Helmholtz equation in spherical coordinates then results in

$$\frac{1}{r^2}\frac{\partial^2 r^2 p(\mathbf{r})}{\partial r^2} + \frac{1}{r^2 \sin^2(\vartheta)}\frac{\partial^2 p(\mathbf{r})}{\partial \varphi^2} + \frac{1}{r^2 \sin\vartheta}\frac{\partial}{\partial \vartheta}\left(\sin\vartheta \frac{\partial p(\mathbf{r})}{\partial \vartheta}\right) + k^2 p(\mathbf{r}) = 0, \tag{4.3}$$

which is a homogeneous second order partial differential equation (PDE).

The Helmholtz equation (4.3) is solved using a product ansatz [Wil99]

$$p(r, \varphi, \vartheta) = R(r)\Phi(\varphi)\Theta(\vartheta), \tag{4.4}$$

that separates the PDE into three ordinary differential equations (ODE) that can be solved seperately:

$$\frac{\partial^2 \Phi}{\partial \varphi^2} + \mu^2 \Phi = 0, \tag{4.5}$$

$$\frac{1}{\sin\vartheta}\frac{\partial}{\partial \vartheta}\left(\sin\vartheta \frac{\partial \Theta}{\partial \vartheta}\right) + \left[\nu(\nu+1) - \frac{\mu^2}{\sin^2\vartheta}\right]\Theta = 0, \tag{4.6}$$

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial R}{\partial r}\right) + k^2 R - \frac{\nu(\nu+1)}{r^2} R = 0, \tag{4.7}$$

where μ and ν are arbitrary constants of non-integer value. The particle velocity is proportional to the derivative of the pressure functions and vanishes in normal direction to the boundary. Therefore sound-hard boundaries (particle velocity = 0) imply Neumann boundary conditions.

4.1 Angular Solutions for the chosen Geometry

Sound-hard boundary surfaces force the perpendicular particle velocity to vanish. Further it is proportional to the derivative of the pressure functions. Physical angular solutions can therefore be found using two-point boundary conditions for the spatial derivative of the sound pressure, so called Neumann boundary conditions.

4.1.1 Admissible Azimuth Solution Under Sound-hard Boundary Conditions

Equation (4.5) is a homogeneous second order ODE with constant coefficients. The solution for the lossless case is [Wil99]

$$\Phi(\varphi) = \Phi_1 \sin(\mu\varphi) + \Phi_2 \cos(\mu\varphi), \tag{4.8}$$

with Φ_1 and Φ_2 being complex valued constants.
The azimuthal Neumann boundary conditions at $0 \leq \varphi_1 \leq \varphi_2 \leq 2\pi$ write as

$$\left.\frac{\mathrm{d}\Phi_\mu(\varphi)}{\mathrm{d}\varphi}\right|_{\varphi=\varphi_{1,2}} = 0. \tag{4.9}$$

They select as admissible solutions [PP14]

$$\Phi_{\mu_m}(\varphi) = N_{\mu_m} \cos\left(\mu_m \varphi\right), \tag{4.10}$$

where the *degree* μ_m is

$$\mu_m = \frac{\pi m}{|\varphi_2 - \varphi_1|} \text{ for } m = 0, 1, 2, \ldots, \infty. \tag{4.11}$$

With the normalization factor N_{μ_m}

$$N_{\mu_m} = \begin{cases} \sqrt{\dfrac{2}{|\varphi_2 - \varphi_1|}} & \mu_m \neq 0 \\ \sqrt{\dfrac{4}{|\varphi_2 - \varphi_1|}} & \mu_m = 0 \end{cases} \tag{4.12}$$

the azimuth solutions forms an orthonormal basis for $\varphi \in [\varphi_1, \varphi_2]$:

$$\int_{\varphi_1}^{\varphi_2} \Phi_{\mu_m}(\varphi)\, \Phi_{\mu_{m'}}(\varphi)\, d\varphi = \delta_{m\,m'}. \tag{4.13}$$

Note:
Generally, the degree μ_m is of non-integer value. With the chosen azimuthal opening angle of $\varphi = 90°$, eq. 4.11 yields values for $\mu_m = 0, 2, 4 \ldots \infty$ (see fig. 13). It can easily be understood that only the $\cos$ part of the general azimuth solution 4.8 has to be considered: the argument of the $\cos$ is scaled such that its extrema are at $\varphi_{1,2}$ (and thus, its derivative vanishes).

4.1.2 Admissible Zenith Solution Under Sound-hard Boundary Conditions

Following [Wil99, PZ13] the so-called *associated Legendre differential equation*, eq. (4.6), is solved for $cos(\vartheta)$ and $0 \leq \vartheta_1 \leq \vartheta_2 \leq \pi$:

$$\Theta(\vartheta) = \Theta_1 P_{\nu_{nm}}^{\mu_m}(\cos(\vartheta)) + \Theta_2 Q_{\nu_{nm}}^{\mu_m}(\cos(\vartheta)), \tag{4.14}$$

where Θ_1 and Θ_2 are in general complex valued constants, $P_{\nu_{nm}}^{\mu_m}$ and $Q_{\nu_{nm}}^{\mu_m}$ denote the *associated Legendre functions* of first and second kind, respectively with *order* ν_{nm} and *degree* μ_m.
For the physical solution we combine the two functions with two scalar weighting factors α_{nm} and β_{nm} [PZ13].

$$\Theta_{\nu_{nm}}^{\mu_m}(\vartheta) = \alpha_{nm}\, P_{\nu_{nm}}^{\mu_m}(\cos\vartheta) + \beta_{nm}\, Q_{\nu_{nm}}^{\mu_m}(\cos\vartheta). \tag{4.15}$$

Neumann boundary conditions are introduced at ϑ_1 and ϑ_2, that also exclude the singularities at the north and south pole

$$\begin{aligned} \left.\frac{\mathrm{d}\Theta_{\nu_{nm}}^{\mu_m}(\vartheta)}{\mathrm{d}\vartheta}\right|_{\vartheta=\vartheta_{1,2}} &= 0 \\ \left[\alpha_{nm}\frac{\partial P_{\nu_{nm}}^{\mu_m}(\cos\vartheta)}{\partial\vartheta} + \beta_{nm}\frac{\partial Q_{\nu_{nm}}^{\mu_m}(\cos\vartheta)}{\partial\vartheta}\right]_{\vartheta=\vartheta_{1,2}} &= 0. \end{aligned} \tag{4.16}$$

A set of μ_m is given from the physical azimuth solution (see chapter 4.1.1). For every μ_m the corresponding ν_{nm} fulfilling the Neumann boundary conditions in zenith are calculated using a zero-finding algorithm. The calculation of Legendre functions of non-integer order was done with an algorithm presented in [OS83]. From the infinite, discrete set of solutions we are only interested in solutions ν_{nm} up to a certain maximum order N. The zero-finding algorithm delivers orthogonal solutions, but α_{nm} and β_{nm} contain an arbitrary common scaling factor. Orthonormality is achieved by suitable scaling of α_{nm} and β_{nm}. A normalization for the zenith solution is introduced so that the following relation holds

$$\int_{\vartheta_1}^{\vartheta_2} \Theta_{\nu_{nm}}^{\mu_m}(\vartheta)\, \Theta_{\nu_{n'm'}}^{\mu_{m'}}(\vartheta)\, d\vartheta = \delta_{nn'}\, \delta_{mm'}. \tag{4.17}$$

Fig. 12 shows the evaluation of eq. (4.15) with α_{nm} and β_{nm}. The corresponding set of μ

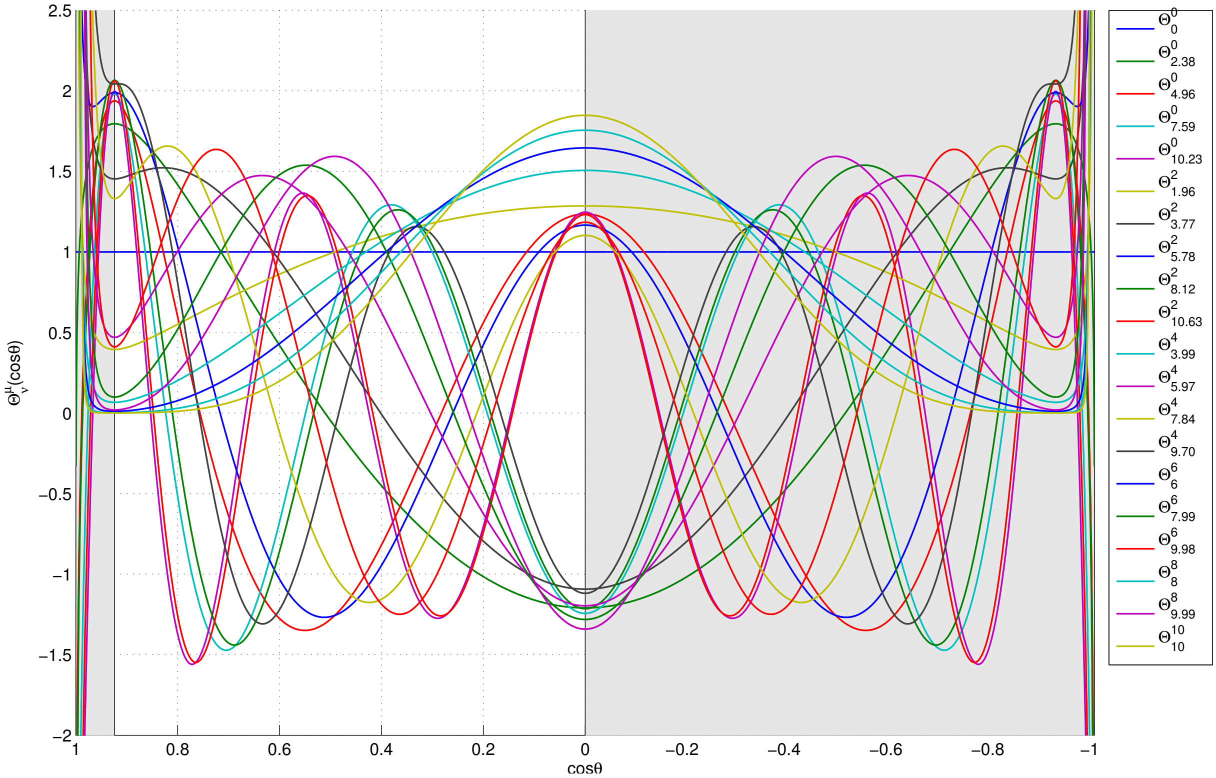


Figure 12: A set of 20 zenith solutions $\Theta_{\nu_{nm}}^{\mu_m}$ for μ and ν and the limiting order of $N = 11$. Note the derivative vanishes at $\cos\vartheta = 0$ and 0.9239 due to zenith Neumann boundary conditions at $\vartheta = 22.5°, 90°$.

and ν in fig. 13 fulfil the boundary design $\varphi = 0°, 90°$ and $\vartheta = 22.5°, 90°$ (see chapter 3).

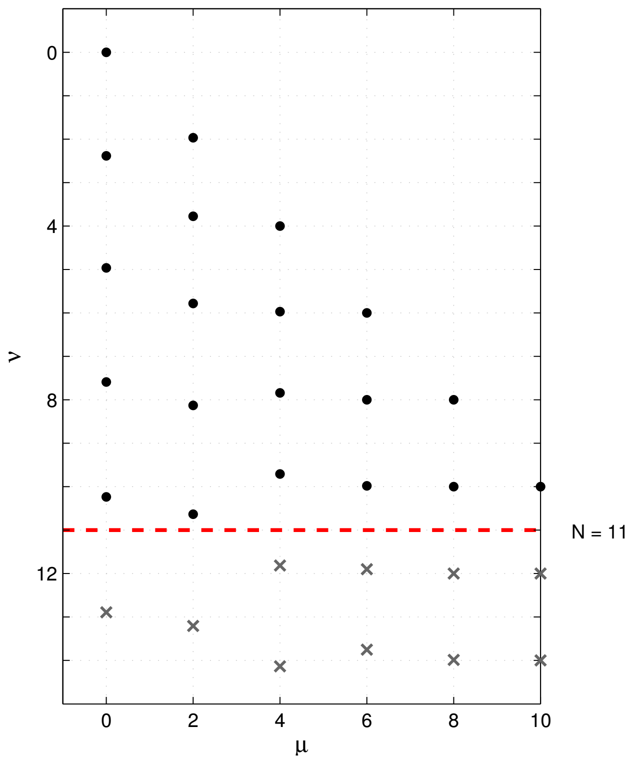


Figure 13: Set of 20 values for μ and ν for the Neumann boundary conditions at $\varphi = 0°, 90°$ and $\vartheta = 22.5°, 90°$ and the limiting order of $N = 11$

4.1.3 Spherical Sector Harmonics

Analogously to spherical harmonics (SH), we introduce the term spherical sector harmonics (SSH) for the combination of the Helmholtz equation's angular solutions fulfilling the two-point boundary conditions in ϑ and φ

$$Y_n^m(\varphi, \vartheta) = \Theta_{\nu_{nm}}^{\mu_m}(\vartheta)\, \Phi_{\mu_m}(\varphi). \tag{4.18}$$

According to the definitions above, the SSH form a complete orthonormal basis on the spherical sector S^2 ($\varphi_1 \leq \varphi \leq \varphi_2$ and $\vartheta_1 \leq \vartheta \leq \vartheta_2$). The following equation holds

$$\int_{\varphi_1}^{\varphi_2} \int_{\vartheta_1}^{\vartheta_2} Y_n^m(\varphi, \vartheta)\, Y_{n'}^{m'}(\varphi, \vartheta)\ \sin(\vartheta)\ \mathrm{d}\vartheta\ \mathrm{d}\varphi = \delta_{nn'}\, \delta_{mm'}. \tag{4.19}$$

The spherical sector harmonics Y_n^m are depicted in fig. 14.

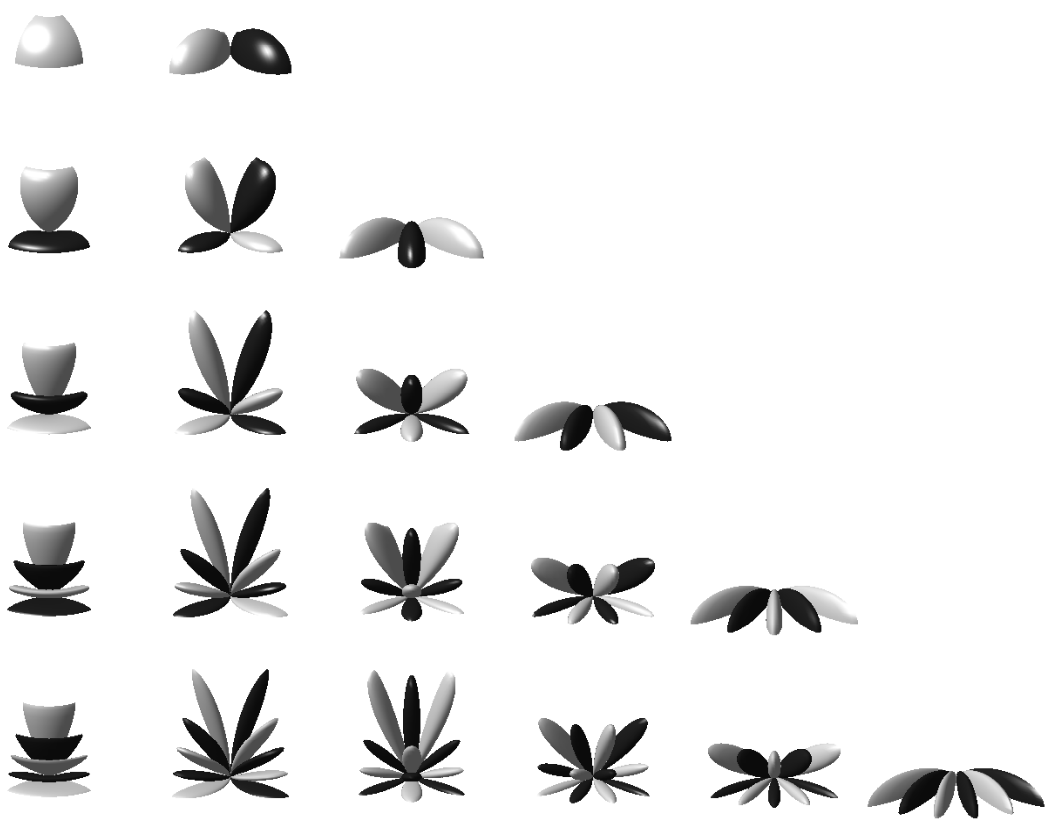


Figure 14: Set of 20 SSH for the values of μ and ν calculated for the Neumann boundary conditions at $\varphi = 0^\circ, 90^\circ$ and $\vartheta = 22.5^\circ, 90^\circ$ and the limiting order of $N = 11$

4.2 Radial Solution

The sound projector is considered an exterior problem under the simplifying assumption of infinite-length boundaries. This means all the sources are located within a finite radius. As a consequence, only the spherical Hankel functions of second kind $h_n^{(2)}$ are solutions to the radial differential equation (eq. 4.7). The pressure at projection radius r is [Wil99]

$$p(r,\varphi,\vartheta) = \sum_n \sum_m c_{nm}\, h_{\nu_{nm}}^{(2)}(kr)\, Y_n^m(\varphi,\vartheta), \tag{4.20}$$

where c_{nm} denotes the *wave spectrum* that is independent of any radius. For large projection radii the far field approximation for $h_\nu^{(2)}$ can be used [PB09]

$$\lim_{x\to\infty} h_\nu^{(2)}(x) = j^{\nu+1}\,\frac{e^{-jx}}{x}, \tag{4.21}$$

simplifying eq. (4.20). The linear phase term e^{-jx} is omitted because it just corresponds to a time delay.

$$\lim_{r\to\infty} p(r,\varphi,\vartheta) = \sum_n \sum_m c_{nm}\, j^{\nu_{nm}+1}\,\frac{e^{-jkr}}{kr}\, Y_n^m(\varphi,\vartheta). \tag{4.22}$$

The particle velocity is controlled by loudspeakers at the surface of the array at $r = a$, which defines c_{nm}. In SSH domain it is referred to as the *spherical wave spectrum* of the particle velocity ζ_{nm} is

$$\zeta_{nm} = \frac{j}{\rho_0 c} c_{nm}\, h_{\nu_{nm}}^{\prime(2)}(ka)\, Y_n^m(\varphi,\vartheta). \tag{4.23}$$

Note: The spherical wave spectrum ζ_{nm} depends on the radius. Because of imposed particle velocity at a, the derivative $h_{\nu_{nm}}^{\prime(2)}(ka)$ of the Hankel function of second kind has to be used.

So called radial steering filters $H_{nm}(ka)$ are used to control the particle velocity in order to achieve the desired directivity pattern in the far field ψ_{nm}. (Thus, they compensate for the frequency-dependent radiation term.)
The *radial steering filters* use the far field approximation (eq. (4.21)) to calculate sound pressure at r from the particle velocity at a. They write as

$$H_{nm}(ka) = \frac{\rho_0 c}{k}\,\frac{j^\nu}{h_{\nu_{nm}}^{\prime(2)}(ka)}. \tag{4.24}$$

Evaluating eq. (4.24) over frequency (with $a = 235\ \mathrm{mm}$, for the various orders ν) results in the necessary filter weights for radial steering. Towards low frequencies the gain for higher orders exceeds the usual gain limit of $40\ \mathrm{dB}$.

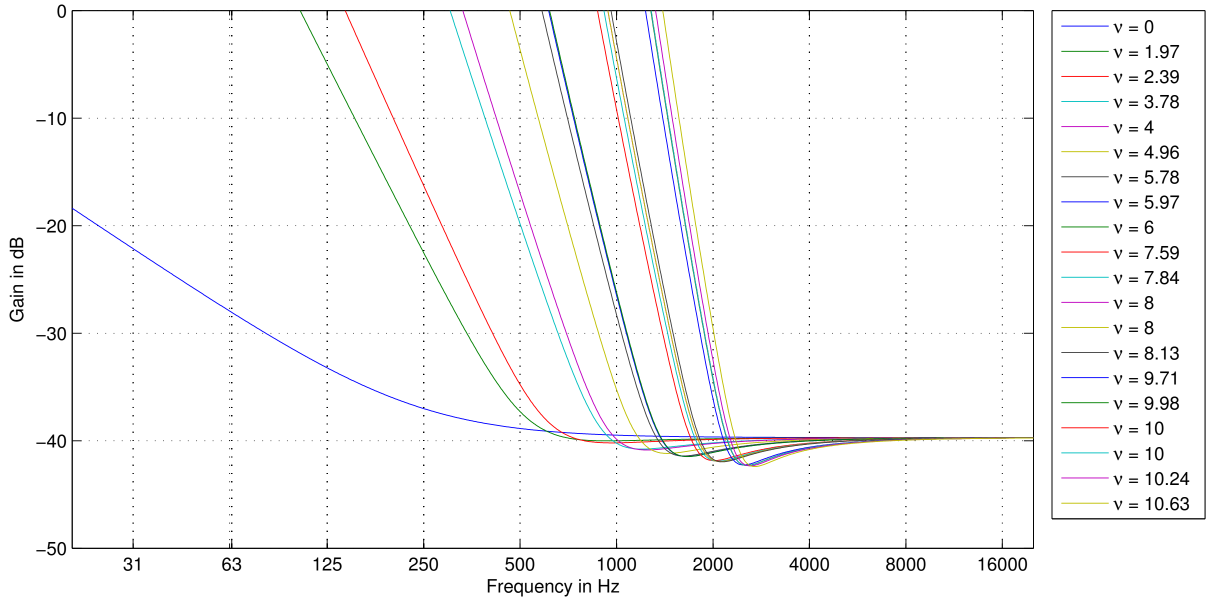


Figure 15: Radial filters for orders ν used for far field extrapolation, $a = 235mm$

4.3 Loudspeaker Aperture

For simplicity, driving the array's loudspeakers can be modeled as controlling the membrane velocities. Following [Zot09], cap-shaped aperture functions $a_l(\varphi, \vartheta)$ are introduced:

$$a_l = \begin{cases} 1 & \text{for points inside l}^{\text{th}} \text{ membrane area} \\ 0 & \text{elsewhere} \end{cases} \tag{4.25}$$

The corresponding surface particle velocity $z(\varphi, \vartheta)$ is described by superimposing the aperture functions $a_l(\varphi, \vartheta)$

$$z(\varphi, \vartheta) = \sum_{l=1}^{20} a_l(\varphi, \vartheta)\, v_l, \tag{4.26}$$

where v_l denotes the membrane velocity. To get the surface velocity in the SSH domain, the equation is transformed. To this end, the 20 aperture functions $a_l(\varphi, \vartheta)$ and the 20 SSH $\boldsymbol{Y}_{nm}(\varphi, \vartheta)$ that should be controlled are sampled by a grid of 22500 HEALPix points [GHB+05], yielding the 22500×20 matrices $\boldsymbol{A}$ and $\boldsymbol{Y}_{grid}$. Expressed discretized, Eq. (4.26) writes as

$$\boldsymbol{z} = \boldsymbol{A}\,\boldsymbol{v}. \tag{4.27}$$

As matrix equations, the SSH expansion $z = \sum_{nm} Y_{nm}\, \zeta_{nm}$ is expressed by

$$\boldsymbol{z} = \boldsymbol{Y}_{grid}\, \boldsymbol{\zeta}. \tag{4.28}$$

The particle velocity in SSH then can be written as a combination of eq. 4.27 and eq. 4.28

$$\boldsymbol{\zeta} = \boldsymbol{Y}_{grid}^{-1} \boldsymbol{A}\, \boldsymbol{v}. \tag{4.29}$$

All spherical cap functions $\boldsymbol{Y}_{cap}$, transforming 20 loudspeaker areas to 20 SSH, yield

$$\boldsymbol{Y}_{cap} = \boldsymbol{Y}_{grid}^{-1}\, \boldsymbol{A}. \tag{4.30}$$

The sampled cap functions are visualized in figure 16.

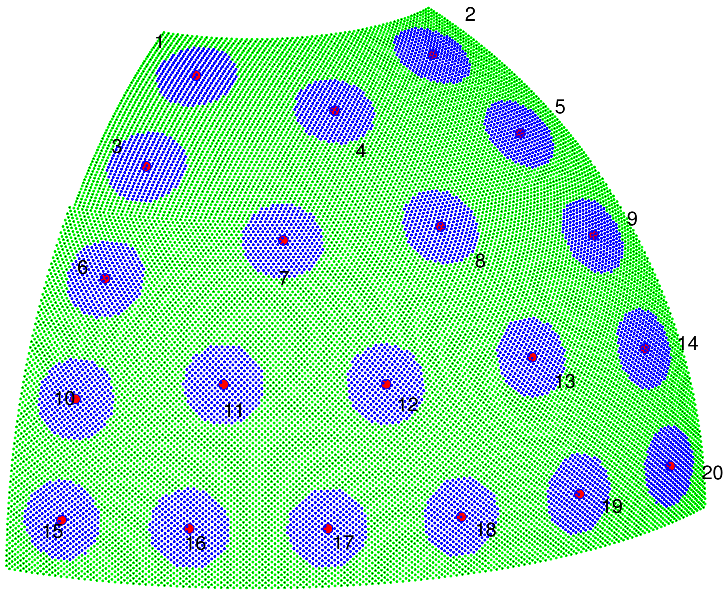


Figure 16: Spherical Sector with all the sampled loudspeaker apertures $\boldsymbol{A}$ membranes are depicted in blue, loudspeaker center position are depicted in red

To speed up the search for an optimal loudspeaker arrangement, we assume that optimization yields useful results also with loudspeakers' cap apertures simplified to a set of spherical Dirac δ distributions at the cap centers.

Further investigations on the loudspeaker aperture are presented in appendix B.

5 Finding an Optimal Arrangement of Loudspeakers

It is never possible to control a continuum of the particle velocity with loudspeakers of a certain size. A limited number of loudspeakers leads to a limited set of SSH modes that can be used to control radiation patterns. However, sound radiation strongly attenuates higher order SSH modes. Thus, they contribute little to radiation patterns. Realistic control of particle velocity is limited in order and discrete in position. The problem of control can be solved with a limited number of loudspeaker positions. As this facilitates accurate control of the surface vibration modes within the orders below $N = 11$, the loudspeakers exciting the surface velocity need to well-distributed. Mathematically, the matrix $\boldsymbol{Y}_{cap}$ (containing the SSH coefficients of well-distributed loudspeakers for orders below $N = 11$) needs to be easily invertible, i.e. well-conditioned, for control. This means the optimization criterion is the condition number of the matrix with respect to inversion. The constraints for the positioning are the following:

- Size of the loudspeaker: loudspeakers cannot overlap and need to have a certain minimum distance due to the firmness of the spherical sector
- Distance to bounding surfaces: because of the cabinet, loudspeakers cannot be positioned too close to the bounding surfaces

When representing the loudspeaker membranes by a spatial Dirac δ distribution at their center (see chapter 4.3), the search for their optimal positions is the same task as with microphone arrays. There, algorithms searching for the optimum on a discrete grid had an advantage over continuous, non-linear optimization. Therefore, the initial choice was to implement a search routine for a minimum condition number on a discrete grid only.

5.1 Simple Iterative Search

Thomas Kößler [Köß11] and Florian Pausch [Pau13] suggest a procedure based on a fine grid of HEALPix points. An initial set of N points is selected randomly. In each iteration, a single point is moved to all vacant positions on the grid, while keeping all the other $N-1$ positions constant. The position that yields the lowest condition number of the inverted transformation matrix is kept. One set of successive iterations for the position of the N sampling nodes defines a run. If the condition number is improved, the set of points will be used as the initial distribution for the next run. This can be viewed as an optimizer with discrete steps, just optimizing a single position at a time. Because the problem actually has N dimensions to optimize, this procedure does not necessarily find the minimum directly, even if the problem is convex. This approach is quite efficient in delivering a well-distributed set of nodes without getting stuck in local minima. Finding a set of points with a condition number $\kappa < 3$ takes computation time of about an hour.

Because of this, a few more sophisticated algorithms were implemented to achieve lower condition numbers at even a lower computational cost.

5.2 Fast Subset Selection

It is a standard procedure in mathematics to approximate a system by considering a wisely chosen subset. Two promising algorithms, both presented in [AB13], are examined in the following. In both cases, the sector is covered with a fine grid of HEALPix points [GHB+05]. The matrix transforming these points to 20 SSH is a short-and-fat matrix.

- **Selection by Greedy Removal:** In each step the column whose deletion yields the lowest condition number is deleted until we end with a square matrix.
- **Greedy Selection:** Starting with one suitable column, in each step the one column that yields the lowest condition number is added, thus reaching a square matrix.

Fig. 17 shows the condition number progression. Both algorithms are very fast, but they are not well suited for the problem of critical sampling. In the last few iterations the condition number almost explodes. The algorithms are ideal to cancel or to add a single column at a time, however, as they do not consider the desired number of columns, they are suboptimal.

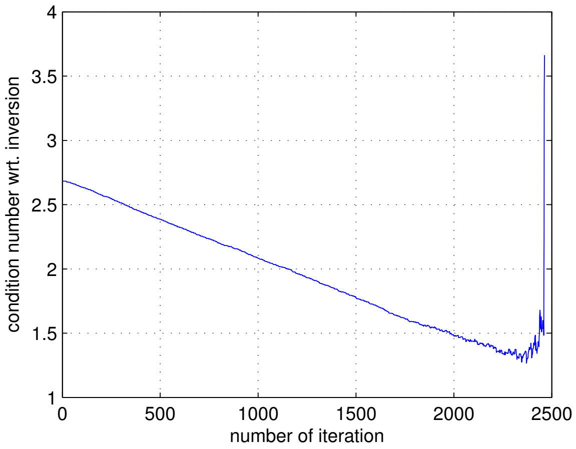


Figure 17: Condition number κ over iteration steps with (left) Greedy Removal algorithm, (right) Greedy Selection algorithm

5.3 Nonlinar condition number minimization (non-grid based)

MATLAB's Optimization Toolbox can be used to optimize the coordinates of the loudspeaker positions with the condition number $\kappa(\boldsymbol{Y}_\delta)$ as the minimization criterion. However, these optimizers are not determined to find the global minimum. The crucial point is the initial set of coordinates. They determine the optimization result.
As a first try the coordinates provided by the *Greedy Selection* algorithm were fed into MATLAB's *unconstrained nonlinear least-squares algorithm*. It simultaneously optimizes the coordinates φ and ϑ of the 20 loudspeaker positions w.r.t. condition number $\kappa(\boldsymbol{Y}_\delta)$. The loudspeaker's minimum distance to each other as well as to the border regions were introduced as a strong penalty for κ. The result was a nearly vertical-symmetric pattern. Starting with a vertical-symmetric initial set lowered the condition number $\kappa(\boldsymbol{Y}_\delta)$ even further. The optimizer returned a nearly symmetric pattern, which was manually put into perfect symmetry. The chosen loudspeaker positions are depicted in fig. 18.
The condition number of the transformation matrix is as low as $\kappa(\boldsymbol{Y}_\delta) = 1.45$. In comparison to critical sampling of a sphere this value is superb.

With the positions yielded from optimization, the aperture matrix has a condition number of $\kappa(\boldsymbol{Y}_{cap}) = 1.6$. This was taken as proof that the loudspeaker aperture functions can be approximated by a spatial Dirac δ distribution (see chapter 4.3). Further investigations on the Dirac δ approximation of the loudspeaker aperture are presented in appendix B

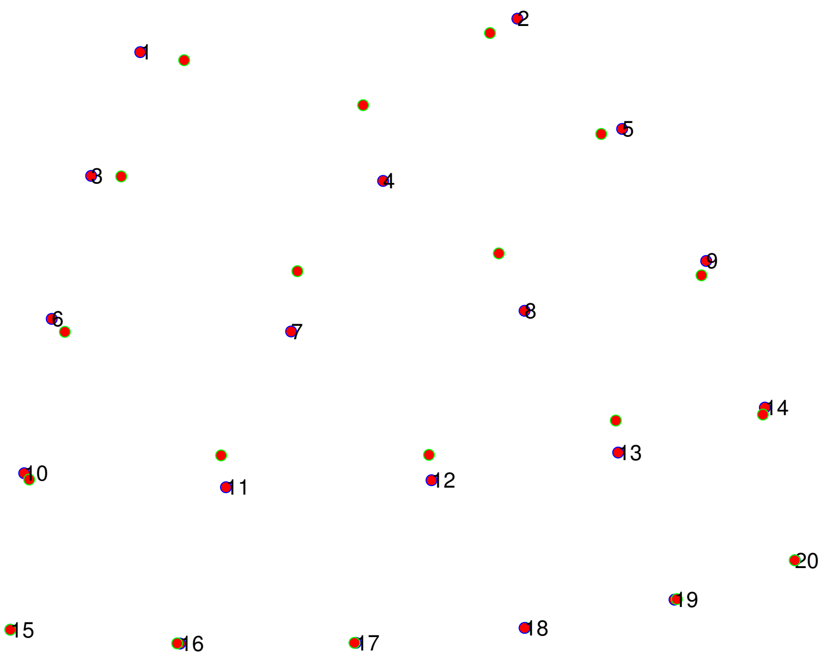


Figure 18: Optimization of the loudspeaker positions
(blue) initial linear spacing, (green) result of the optimization

6 Construction

After finding the optimal positions for the loudspeaker on the spherical sector in MATLAB, the coordinates were transferred to the CAD Files. To double-check that all the loudspeakers fit at their positions, they were modelled and mounted in the CAD environment. A view of the inside of the Sound Projector is depicted in fig. 19.

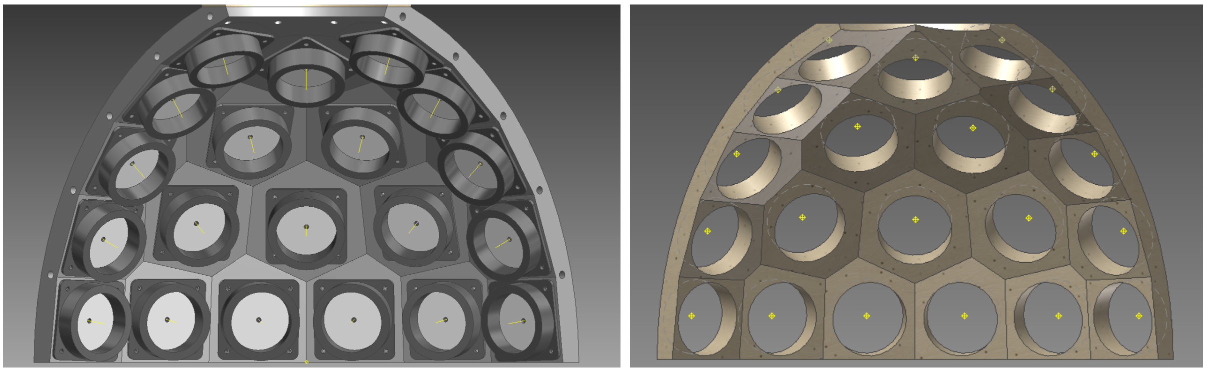


Figure 19: View inside the CAD modelled spherical sector
(left) loudspeakers at their mounting positions, (right) loudspeakers removed

A local carpenter manufactured the cabinet and the spherical sector on the CNC milling machine. After finishing, the joinings of the cabinet were sealed to guarantee air tightness. The loudspeakers were mounted from the backside and fixed with 4 wood screws and metal washers to prevent the plastic to break. Speaker cable is pulled to the top board and soldered to the same type of HARTING Han® connector that is used with the ICO. This way the usage of the same outer cable, connecting it with the sonible d24 amplifier, is possible. A speaker flange was mounted in the bottom board. Before closing the cabinet it was stuffed with Visaton damping material.

7 Measurement

The aim of the conducted measurements was to validate the theory of guided radiation within sound-hard boundaries. Most importantly, they should clarify how well the theoretical assumptions of infinitely long boundaries model the practical results.

7.1 Setup

To examine the radiation characteristics of the sound projector, directional impulse response (DRIR) measurements were made in the anechoic chamber of the IEM. Figure 20 shows the measurement setup.

Figure 20: Setup for IR measurement, sound projector mounted on a stand in the center of the semicircular microphone array

To attain a full-spherical grid, the sound projector was placed on a turntable and positioned in the center of a semicircular microphone array via a stand. Thus an equi-angle grid [DH94, Raf05] of 18 zenith and 36 azimuthal sampling points on a radius of $r = 0.8\,\mathrm{m}$ was gained. The loudspeaker-driving signals were filtered with a 4th-order Butterworth high pass at $100\,\mathrm{Hz}$.

Note:
The room is anechoic for frequencies above $200\,\mathrm{Hz}$. Any results below $200\,\mathrm{Hz}$ can therefore not be evaluated. Furthermore, we have to consider acoustic reflections from the turntable.

7.2 Method

Before measurements, each microphone had to be calibrated using a pistonphone calibrator supplying a sine tone of $1\,\mathrm{kHz}$ at $94\,\mathrm{dB}$. The multiple swept-sine method [Far00, MBL07], with start and stop frequencies of $\omega_1 = 50\,\mathrm{Hz}$ and $\omega_2 = 22050\,\mathrm{Hz}$, used a sweep length of $T = 1.5\,\mathrm{s}$ and an inter sweep interval of $T_I = 1\,\mathrm{s}$ at a sample rate $f_S = 44100\,\mathrm{Hz}$. The impulse responses h are gained via deconvolution by spectral division

$$h = \Re\left\{\mathcal{F}^{-1}\left\{\frac{\mathcal{F}\{y\}}{\mathcal{F}\{s\}}\right\}\right\}, \tag{7.1}$$

where s denotes the sweep signal, y denotes the recorded microphone signal, $\Re$ denotes the real part (to remove numerical inaccuracy beyond 10^{-10}), $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the Fourier transform and its inverse. The IRs are cut to two different lengths:

- 500 samples. This allows to express the low frequency diffraction around the borders but also features possible reflections from the turntable's surface.
- 300 samples. This length of IR does not contain the reflections from the turntable. It is used for the beamforming approach presented in section 7.4.

Towards very low frequencies the measurement shows a low SNR due to the following reasons:

- playback with a protective high pass filter
- low sound pressure of the loudspeaker at low frequencies

- the anechoic room's lower frequency limit for reflections
- low-frequency noise, possible rumbling of passing trams, etc.

Therefore, both sets of IRs are filtered with a second-order butterworth high-pass at 100 Hz. Because it is not intended to use the sound projector below 150 Hz, this filtering does not affect the loudspeakers' frequency response of interest. The IR's onset is faded in with a half Hann window of 50 samples, the decay is faded out with a half Hann window of 100 samples. The cut and faded IRs are depicted in fig. 21.

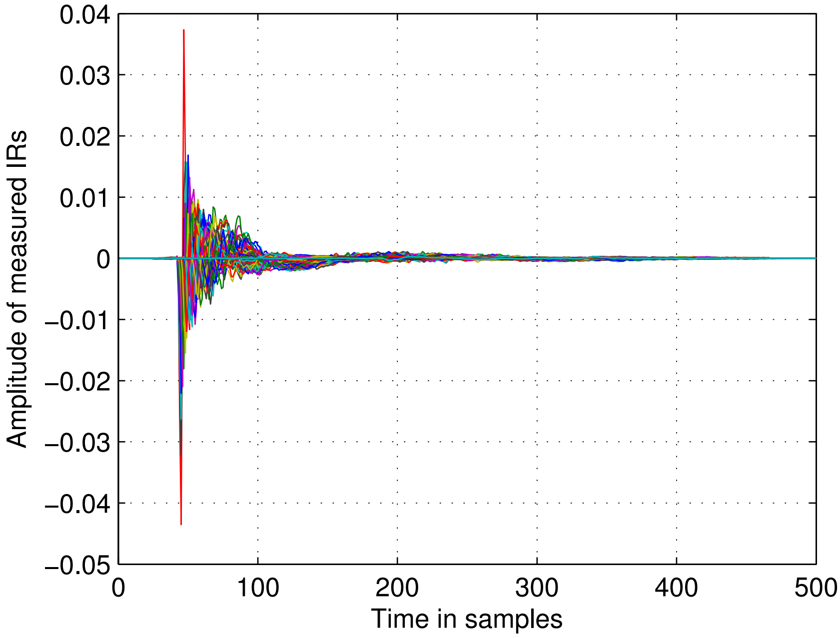


Figure 21: Impulse responses measured from loudspeaker 1 to the grid of 648 microphones

7.3 Results for Individual Loudspeakers

The radiation characteristics of the individual loudspeakers are examined using the measured directional impulse response and evaluating it at a certain frequency. Due to their individual position w.r.t. the bounding surfaces, each loudspeaker shows a different radiation characteristic at the microphone radius. Reflections on the sound hard boundaries lead to constructive and destructive interference for certain frequencies (cf. fig. 22 and 23). Due to these interferences, which are different for every loudspeaker, it appears useless to measure and flatten a loudspeaker's on-axis frequency response. The array is designed to simultaneously drive all loudspeakers. Thus, a certain sound pressure distribution at its spherical sector surface is created, further radiating within the sound-hard boundaries. Therefore, the individual loudspeaker's radiation characteristics do not implicate the capability of the array.

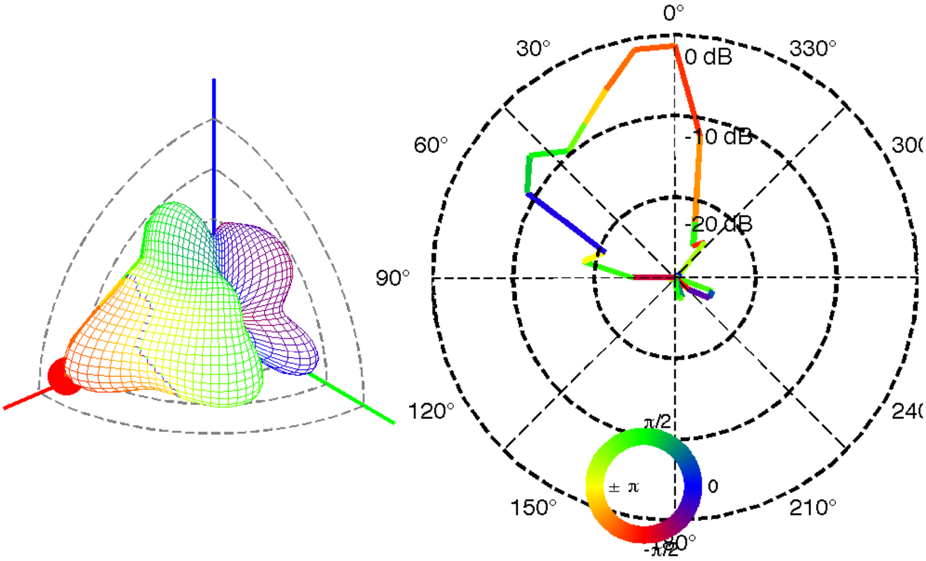


Figure 22: Single loudspeaker's radiation characteristics at $f = 1$ kHz, constructive interference at loudspeaker's position caused by relative position to bounding surfaces

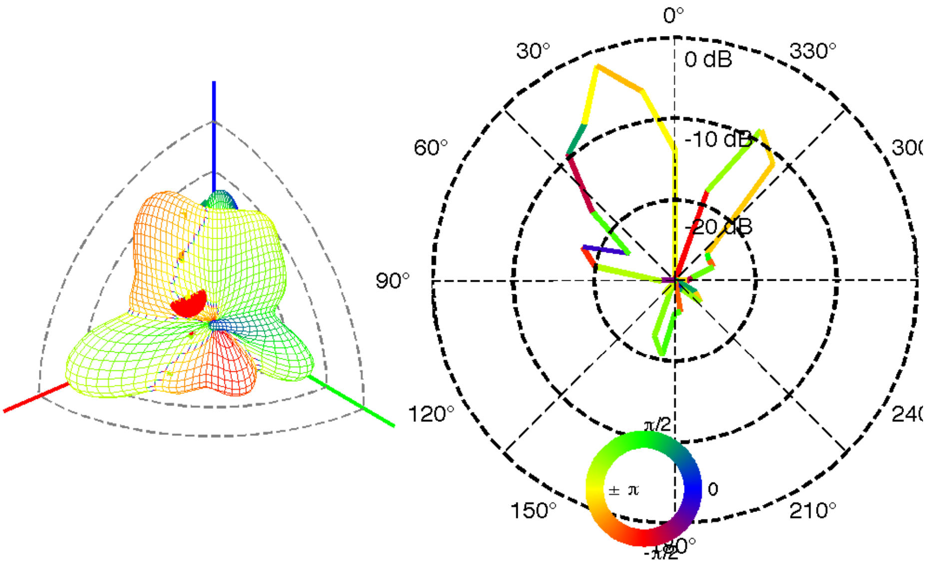


Figure 23: Single loudspeaker's radiation characteristics at $f = 1$ kHz, destructive interference at loudspeaker's position caused by relative position to bounding surfaces

7.4 Beamforming

Loudspeaker beamforming is used to attain a certain pattern of sound radiation. Modal beamforming in SSH utilizes a vector $\boldsymbol{y}_{SSH}$ ($[1 \times 20]$) weighting the 20 SSH coefficients to attain the desired beam characteristic. $\boldsymbol{y}_{SSH}$ describes the steering position in SSH coefficients

$$\boldsymbol{y}_{SSH} = \mathcal{SSH}\{\delta(\vartheta - \vartheta_0, \varphi - \varphi_0)\}, \tag{7.2}$$

where ϑ_0, φ_0 denote the steering direction and $\mathcal{SSH}$ is the spherical sector harmonics transform. The radial filters compensate for frequency dependent attenuation of spherical sector harmonics. The principle is shown in fig. 24.

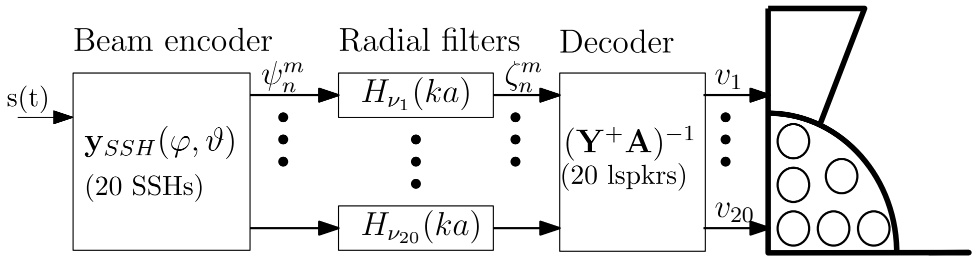


Figure 24: Principle of modal beamforming

However, this is the analytical model for modal beamforming, which is not exactly applicable to our measurement for two reasons:

- Finite length effects and hereby caused diffraction to the projector's rear are best reflected in the surrounding directional sound pressure measurements.
- Such measurements do not provide means to control the particle velocity on the projector's loudspeaker membranes (in contrast to laser vibrometry measurements). Accurate synthesis of directivity pattern requires appropriate near- or farfield equalization [ZN07]. However, in our case we neither have near field on-axis sound pressure nor particle velocity.

The analytical model for modal beamforming shows the basic principle. Encoding of a beam and decoding to particle velocities at the surface are unproblematic. However, the radial filters demand high amplifications. These insights are used in the following section, which describes a method for beamforming based on the measurement data.

7.4.1 Measurement Based Least-Squares Beamforming

Our aim is to calculate a vector for the loudspeaker control $\boldsymbol{u}$. Fig. 25 shows the principle of the modal least squares beamformer derived here.

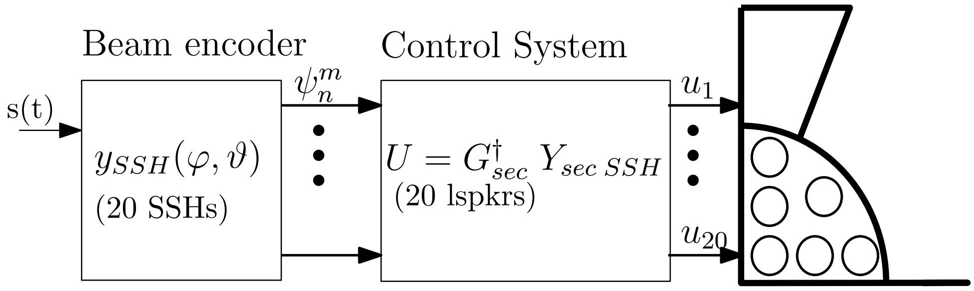


Figure 25: Principle of modal least squares beamforming

The measurement path is stored in the IR matrix $\boldsymbol{P}$ with a length of 500 samples from 20 loudspeakers to 18×36 microphone positions ($[500 \times 20 \times 18 \times 36]$). Calculation is done for a single frequency via a Fourier kernel. The frequency response matrix $\boldsymbol{G}(\omega_k)$ is defined as

$$\boldsymbol{G}(\omega_k) = \mathcal{F}_{\omega_k}\{\boldsymbol{P}\}, \tag{7.3}$$

It has the dimension $[20 \times 18 \times 36]$ (the IR's length / FFT size adds the associated dimension). $\mathcal{F}_{\omega_\mathrm{k}}$ is the Fourier analysis at one frequency ω_k only. The measurement paths with the 70 microphones covering the array's opening sector are selected and stored in $\boldsymbol{G}_{sec}$. With the least-squares method, the following expression has to be minimized

$$\min \|\boldsymbol{G}_{sec}\,\boldsymbol{U} \ - \ \boldsymbol{Y}_{sec\,SSH}\|^2. \tag{7.4}$$

The solution for $\boldsymbol{U}$ is

$$\boldsymbol{U} = \boldsymbol{G}^{\dagger}_{sec}\,\boldsymbol{Y}_{sec\,SSH}, \tag{7.5}$$

where $\boldsymbol{U}$ is a $[20 \times 20]$ matrix describing a linear control system. It converts the 20 desired SSH signals at the position of the microphones to 20 loudspeaker signals. $\boldsymbol{G}^{\dagger}_{sec}$ denotes the Moore-Penrose pseudoinverse of the $[70 \times 20]$ frequency response matrix (20 loudspeakers to the 70 microphones in the sector). $\boldsymbol{Y}_{sec\,SSH}$ is a $[70 \times 20]$ SSH transformation matrix from the 70 microphone positions into 20 SSH. Our goal is to find the loudspeaker control vector $\boldsymbol{u}$ which allows to attain a certain radiation pattern at the microphone position. $\boldsymbol{u}$ ($[20 \times 1]$) is calculated by

$$\boldsymbol{u} = \boldsymbol{U}\,\boldsymbol{y}_{SSH}, \tag{7.6}$$

The impulse response measurements describe the entire radiation from loudspeaker to microphone. The least-squares beamforming approach, yielding the linear control system, contains the loudspeakers' frequency responses, as well as the radiation-dependent attenuation (in the analytical model covered by the radial filters).

Note:
For MATLAB's `pinv(.)` algorithm, a regularization of 0.05 was used. Regularization strengthens the matrix' main diagonal, and thus allows for numerically stable matrix inversion. It enhances prominent components, and attenuates the influence of weak components (which would be amplified strongly with inversion). Thus, it can be interpreted as a way of limiting the effect of most likely negligible components, similarly as the soft knee does with radial filters.

To evaluate the Sound Projector's beamforming capability, the loudspeakers are weighted

$$\boldsymbol{P}_{beam} = \boldsymbol{G}\,\boldsymbol{u}, \tag{7.7}$$

where $\boldsymbol{G}$ is the $[648 \times 20]$ frequency response matrix at a certain frequency. The weighted matrix $\boldsymbol{P}_{beam}$ is further analysed on the 648 microphone position using order $N = 17$ spherical harmonics decomposition $\boldsymbol{Y}_{sh\ p}$

$$\boldsymbol{\psi} = \boldsymbol{Y}_{sh\ p}^{\dagger}\,\boldsymbol{P}_{beam}, \tag{7.8}$$

It also employs interpolation as to be viewed in the following balloon plots. Fig. 26 - 29 show the practicability of the least-squares beamforming approach at a central steering position. A wider beamshape at 300 Hz is visible, also spatial aliasing becomes significant at 4 kHz in fig. 30.

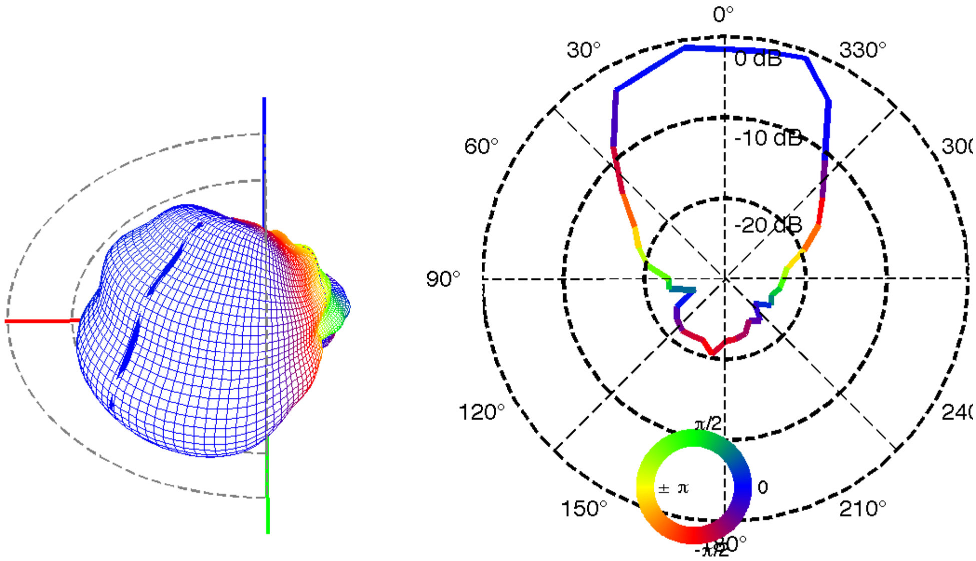


Figure 26: LS beamforming at $f = 300$ Hz,

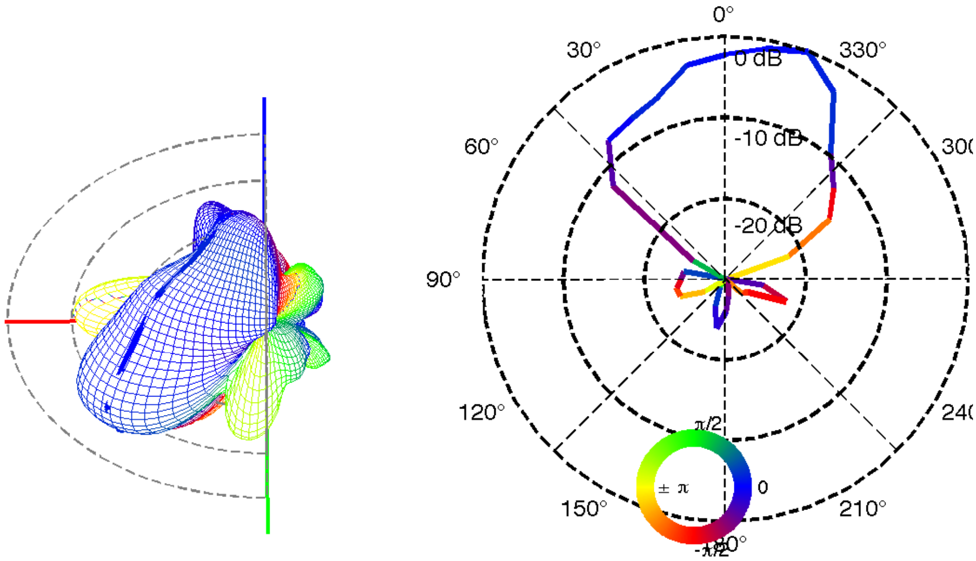


Figure 27: LS beamforming at $f = 500$ Hz,

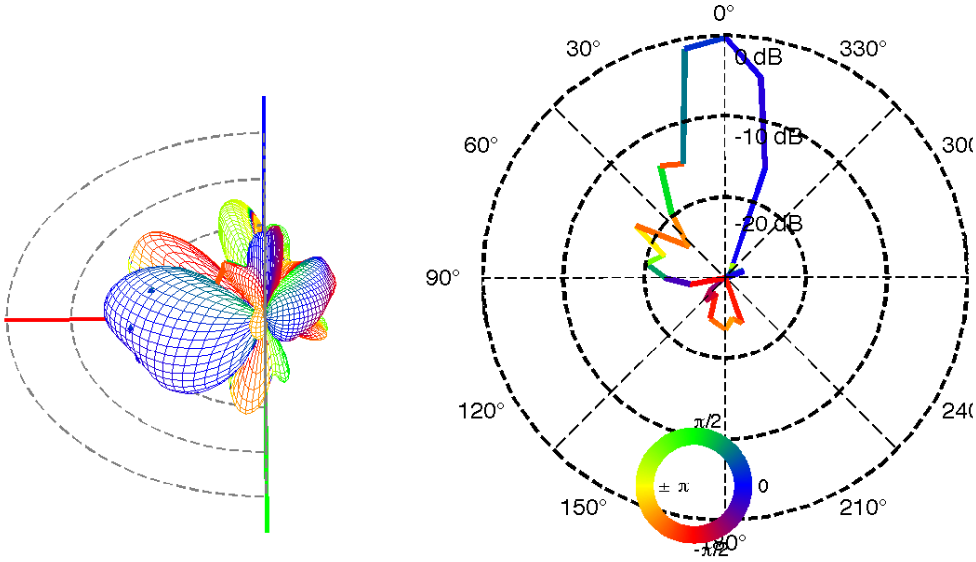


Figure 28: LS beamforming at $f = 1$ kHz,

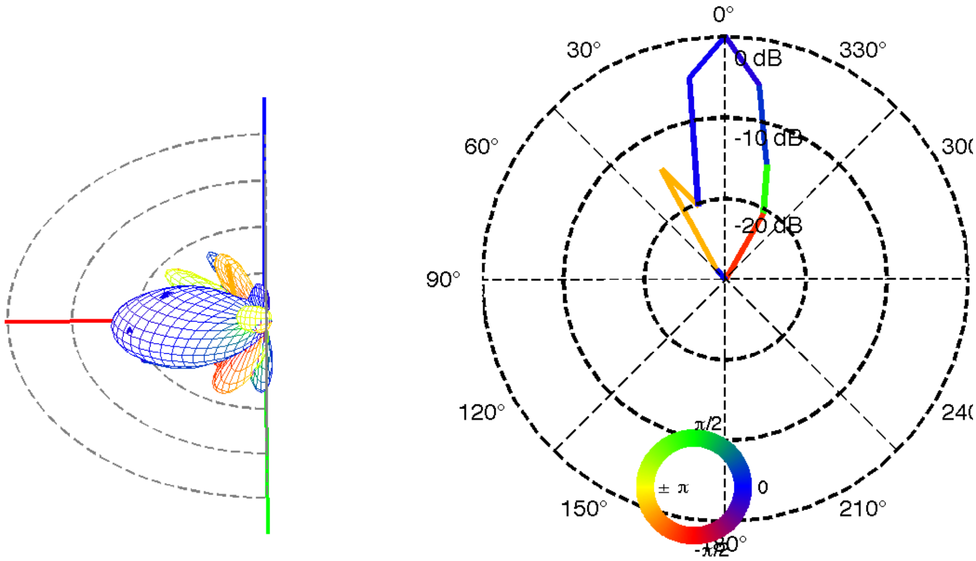


Figure 29: LS beamforming at $f = 3$ kHz,

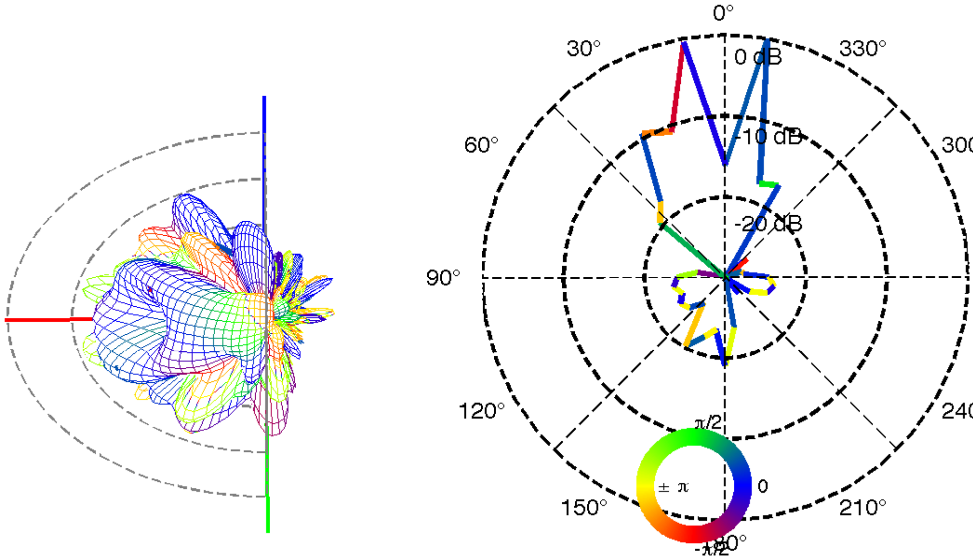


Figure 30: LS beamforming at $f = 4$ kHz, spatial aliasing is clearly detectable

7.4.2 Far Field Extrapolation

The microphones are located just at the border of guided radiation. From there on, any further radiation has to be described in spherical harmonics. To account for the attenuation of low frequencies and higher orders, radial filters for SH have to be considered

$$\psi_r = \psi_{r_0} \frac{h_n(k\,r)}{h_n(k\,r_0)}, \tag{7.9}$$

where r is the radius of extrapolation and r_0 is the radius of the microphone. For a radius of $r = 5$ m the (theoretically necessary) amplification for low frequencies and higher orders is illustrated in fig. 31. If this amplification is not done, the spatial definition and the

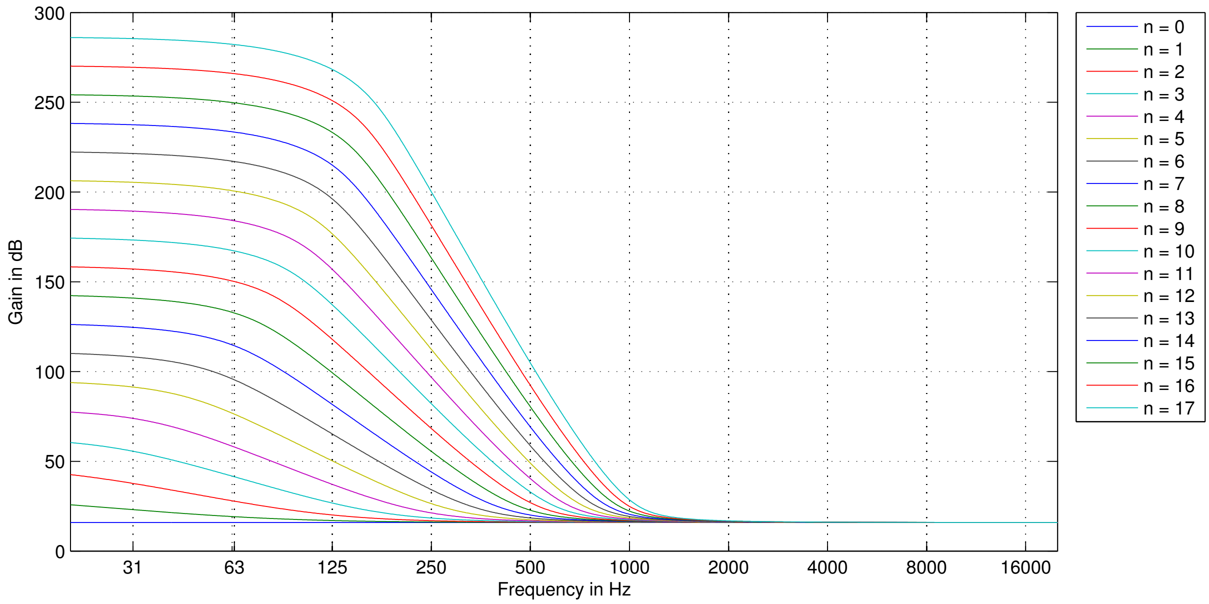


Figure 31: Theoretically necessary radial filter for $r = 5$ m

beam patterns for low frequencies become blurry. This is equivalent to the physical effect of weak spatial defined radiation of low frequencies. Fig. 32 and 33 show the loss of spatial definition beyond the guided radiation at $f = 1$ kHz.

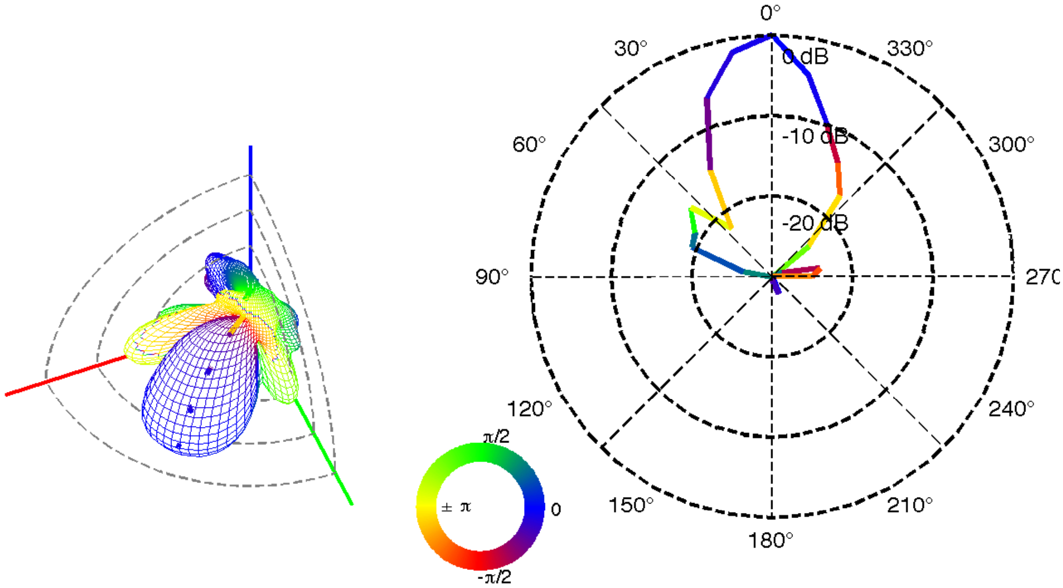


Figure 32: LSE beamforming at $f = 1$ kHz at microphone radius $r_0 = 0.8$ m

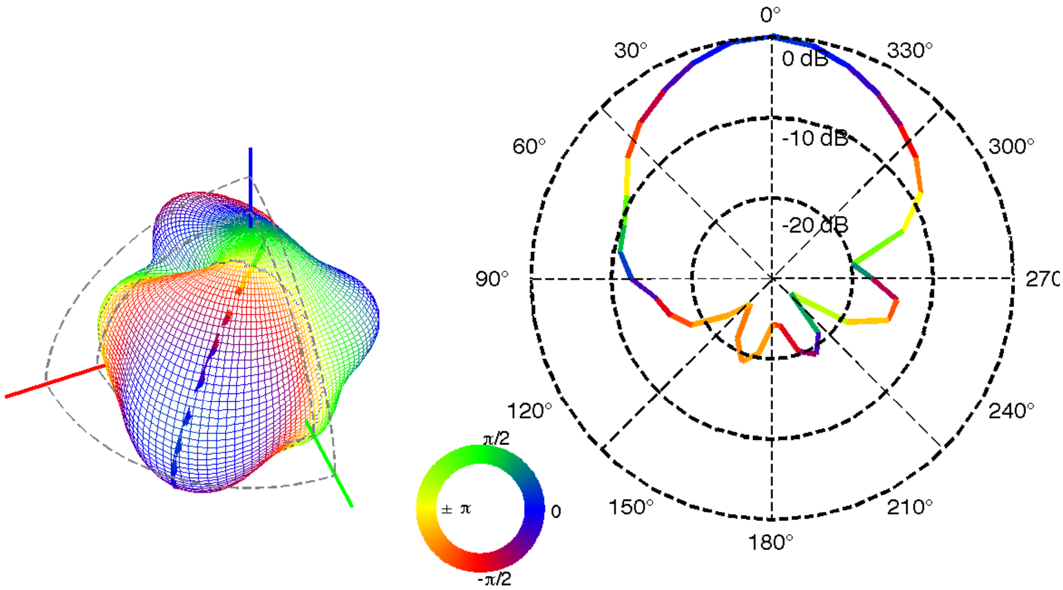


Figure 33: LSE beamforming at $f = 1$ kHz at projection radius $r = 5$ m

7.4.3 Far-field Beamforming Using SSH Radial Filters

The radiation outside of the bounding surfaces actually has to be described in spherical harmonics (SH). Compensation for the loss of spatial resolution for low frequencies can theoretically be obtained via radial filters in the SH domain. However, this implies a transformation from SSH to SH. Alternatively, the whole loudspeaker side could be controlled in SH, which is impracticable (SSH order $N = 11$ leads to 144 channels).

One practicable approach is to apply radial filters H for SSSH , thus amplifying the SSH coefficients directly. Using the far-field approximation they write as

$$H_{nm}(kr_0) = \frac{1}{k}\,\frac{j^{\nu+1}}{h_{\nu_{nm}}(kr_0)}, \tag{7.10}$$

where r_0 denotes the microphone radius. The necessary gain limit of 30 dB can be interpreted as a regularization filter for low frequencies:

$$H_{lim} = \frac{1}{H}\,\frac{|H|}{|H| + 10^{\frac{-30\mathrm{dB}}{20}}}. \tag{7.11}$$

Dependency on n, m and ka was omitted for readability. The limited radial filters H_{lim} are evaluated for r_0 over frequency, as depicted in fig. 34. The SSH radial filters show a slight

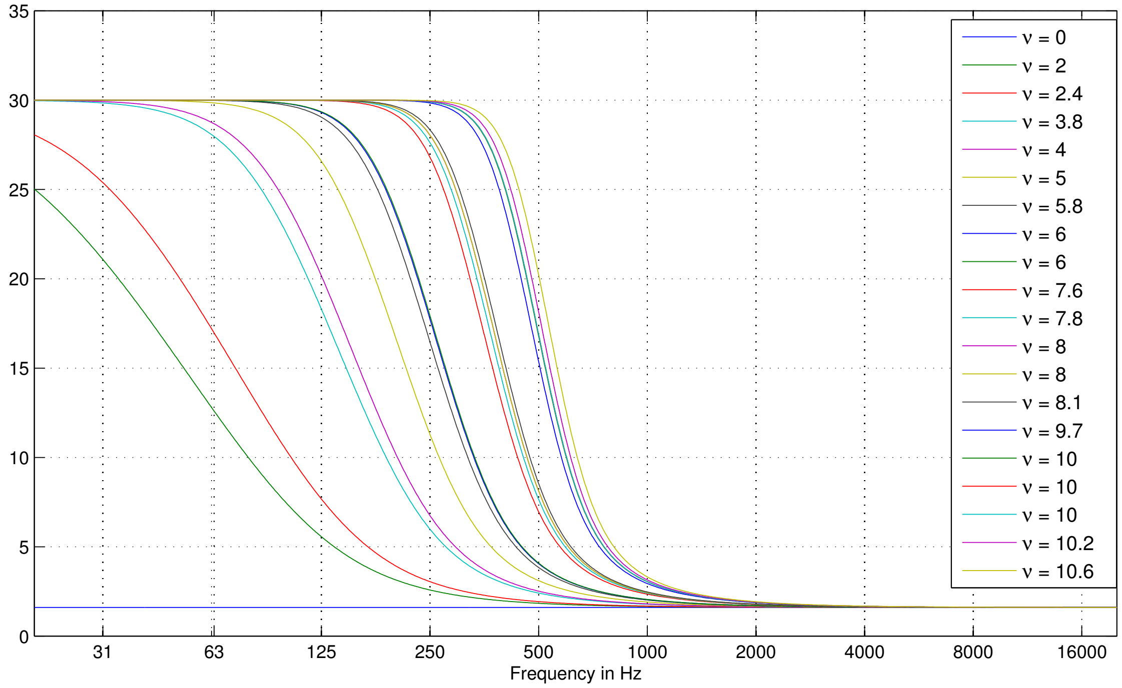


Figure 34: Radial Filters for 20 SSH from $r_0 = 0.8$ m to far-field

improvement in directivity for beamforming, as illustrated in fig. 35 and 36

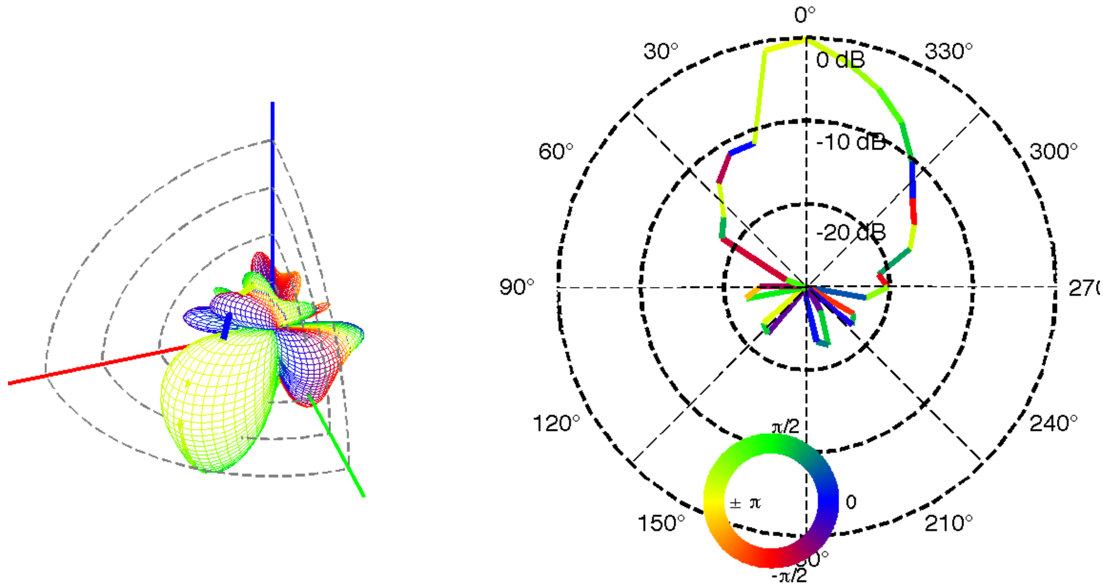


Figure 35: LSE beamforming with radial filters H_{lim} evaluated at $f = 1$ kHz and radius $r = 5$ m

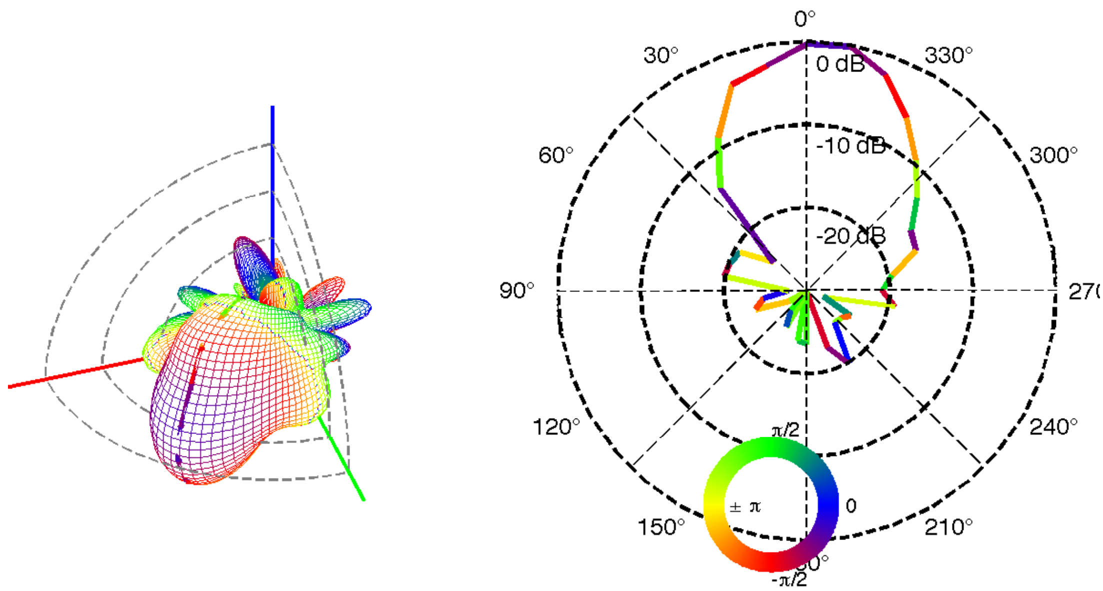


Figure 36: LSE beamforming without radial filters H_{lim} evaluated at $f = 1$ kHz and radius $r = 5$ m

8 Application

8.1 Real-time control interface

A real-time beamforming control interface was implemented in Pure Data (PD). It follows the least-squares approach first implemented in Matlab (see chapter 7.4.1).

For this, the loudspeaker control matrix $\boldsymbol{U}$ (see eq. 7.5) is calculated in the frequency domain and exported as a $[20 \times 20]$ IR matrix of length 2048. It holds the information how to drive the 20 loudspeakers in order to attain a certain radiation pattern at the microphone position (decomposed into 20 SSHs).

In a first step of processing the SSH encoding of the beamsteering is calculated. For this, the normalized combination of legendre functions for the 20 values of ν and μ - evaluated at $22° \leq \vartheta \leq 90°$ in steps of $1°$ - are written to `legendre.mtx`. Also, the cosine function for the 20 values of μ are evaluated at $0° \leq \varphi \leq 90°$ in steps of $1°$ and written to `cos.mtx`. The SSH encoding of the beamsteering is then calculated using these matrices as a lookup table. The principle can be seen on fig. 37 and 38.

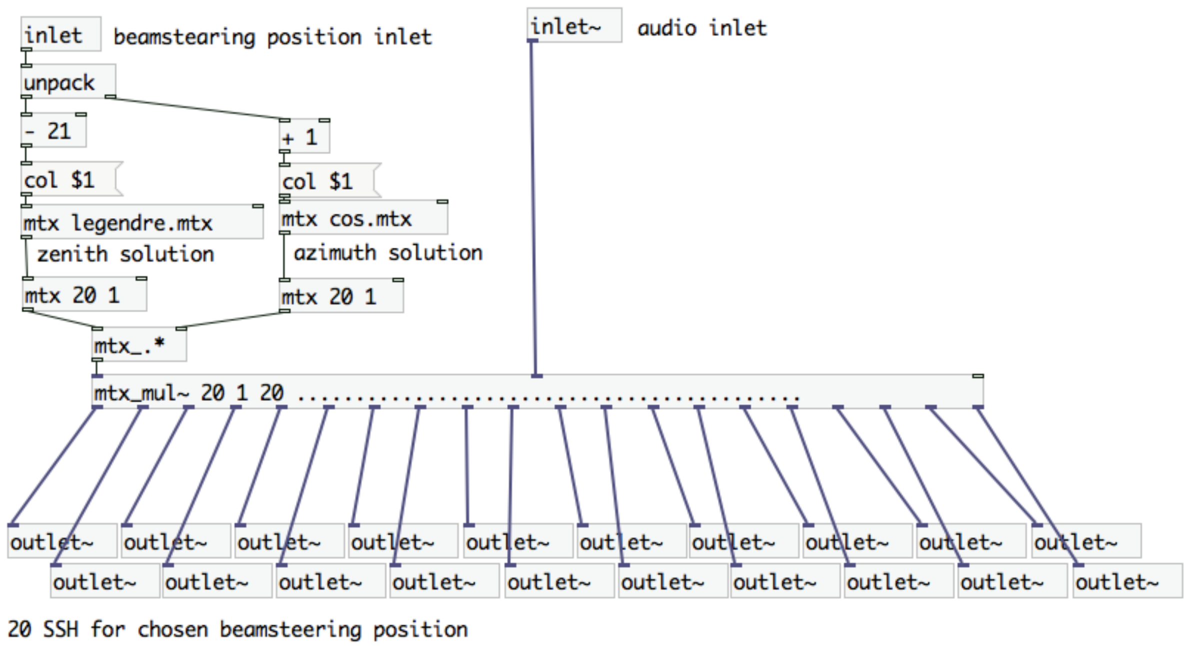


Figure 37: `pd encoder` Beam encoder part of the real-time processing, abstraction in PureData. Zenith solution holds the normalisation as well

The subpatch `pd partconvs` is created automatically by the patch and holds 20×20 convolutions. Furthermore, the patch creates 20×20 tables and reads the loudspeaker

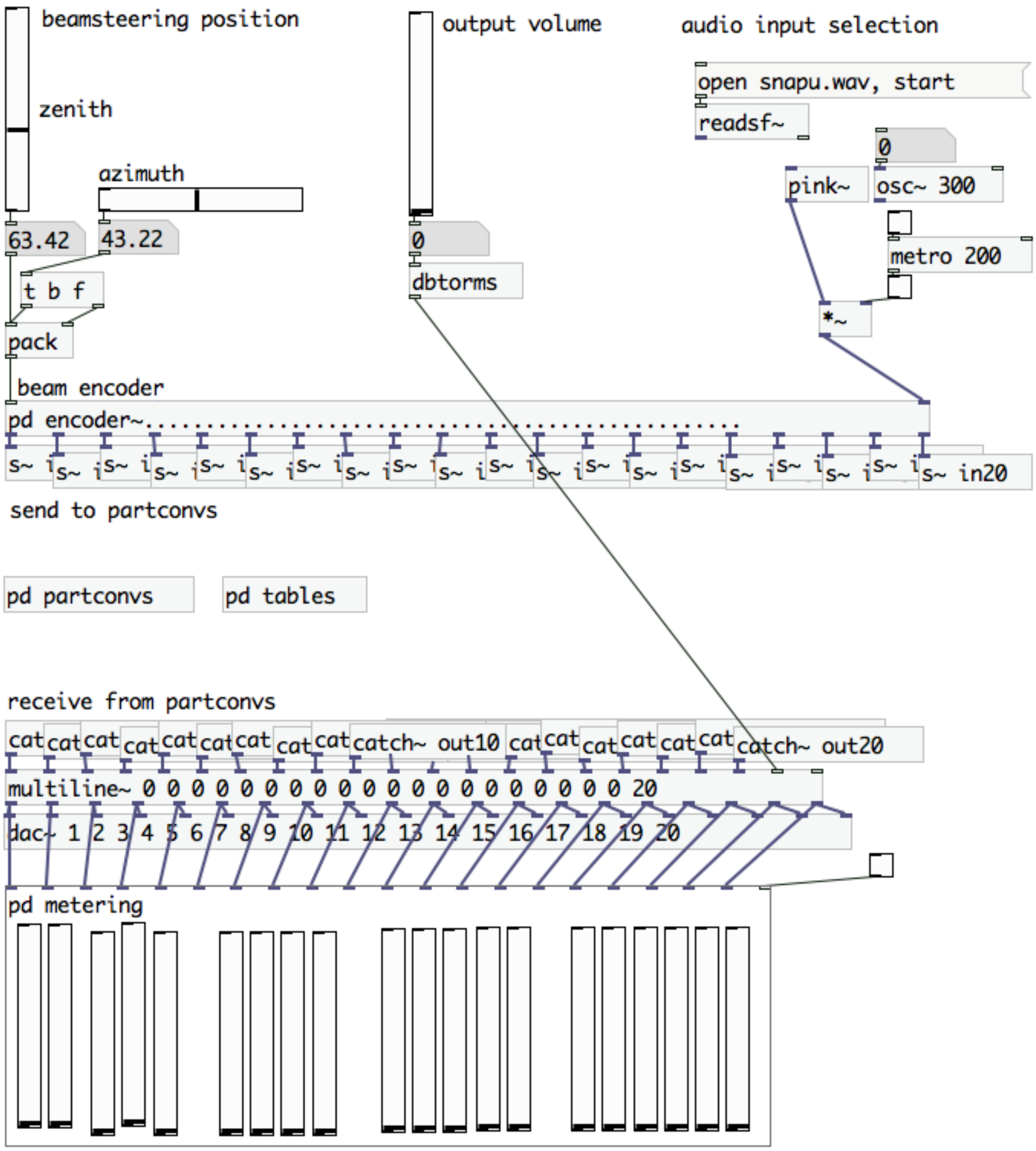


Figure 38: Beamforming interface in PureData

control IR from a .wav file. Processing load is critical but it can be handled on a standard laptop.

8.2 First Impressions

The prototype's capability was examined in the IEM CUBE using the real-time beamforming interface. The array was positioned on the floor facing a wall in a distance of approximately 5 m. To examine the different steering positions, pink noise bursts were used.

The array showed promising beamforming and beamsteering capabilities. Strong wall reflections could be excited. The backward suppression of the array's enclosure is high enough to perceive wall reflections as the louder source. However, at that time, a sufficiently diffuse wall to achieve diffuse reflections was not available. Suitable diffusors are required to orchestrate wall reflections musically and artisticly (see [SZF14]).

Beamsteering close to the border elements showed significantly more high frequencies than in a middle position. Also, there is an ambiguity: due to sidelobes, both extreme positions in azimuth steering sound very similar. However, there are position-dependent spectral variations.

The positioning on the floor is not exactly true to the measurement situation that was used to calculate the loudspeaker control matrix $\mathbf{U}$. In the measurement situation we had almost no reflections from any walls. Here we have early reflections from the floor.

A second position was also examined: To extend the limited bounding surfaces, the prototype was positioned in one corner of the room. This location extends three border elements to the size of the room. Whether this enhances the beamforming capability will have to be examined in the future.

9 Conclusion and Outlook

In this master thesis, a novel sound projector was introduced. The optimal geometry was derived for one operational mode. Spherical Sector Harmonics especially tailored to this geometry are utilized for beamforming.

The measurements of the radiation characteristics reveal a strong acoustic focusing of the sound projector. This is achieved by the guided radiation provided by the sound-hard bounding surfaces in azimuth and zenith. A real-time interface for modal beamforming provides steering of the sound beams. The first impressions show a strong accentuation of acoustic wall reflections.

Finding the Optimal Diffuse Reflector:

The intention of the sound projector was to excite diffuse wall reflections that are perceived as virtual sources. For this, diffuse reflections are needed in a wide range of frequencies. Reflection is a parameter of both, material and surface structure. Diffuse reflection of low frequencies is hard to achieve with a relatively flat wall. It has to be examined in which frequency range diffuse reflections are most important for the perception of virtual sources. Once a suitable diffuse reflector is found, the optimal (fixed) position for the sound projector will facilitate steering to excite distinct reflections.

Mounting the sound projector in the room's upper corners:

Positioning the sound projector in one of the upper corners of a cuboid room extends the effective length of the bounding surfaces up to the room size. So far, the beamforming performance for low frequencies was reduced by the finite length bounding surfaces which allowed diffraction to the backside of the array. Radial filters for far-field beamforming no longer need to compensate for any spherical radiation of low frequencies. Therefore the sound projector would no longer require the enormous bass boost for directed radiation. This effect could already be examined with the sound projector in its existing form. However, there is an optimized geometry for this position: Choosing the zenith and azimuth angle aperture with $90°$ each, simplifies the geometry to an eight-sphere. The necessary three bounding surfaces are supplied by the room's walls.

Analytic Beamforming Approach:

When placing the array in a room's corner, the bounding surfaces extend to the room size. Beamforming then can be performed with the analytical model of chapter 7.4. However, with the analytical model, it becomes necessary to directly control the membrane velocities. The membrane excursion can be measured with the Laser Doppler Vibrometer. Calculating the ratio between driving voltage and membrane excursion yields equalization filters to flatten the frequency response of the loudspeaker. As it was proposed for the ICO, the next step is the implementation of MIMO crosstalk cancellation filters for the sound projector [ZSN08, Lös14].

Further Improvement of the Real-time Control:

The PD patches, which are currently used to steer the beam, work smoothly with the IEM's hardware. However, the ambix plugin suite by Kronlachner [Kro14] is very attractive for numerous reasons: Computational efficiency of the performed convolutions, intuitive control, possible integration in the Reaper software suite, as well as stand-alone capability; Some minor additions should make the ambix plugins work with the SSH domain as well. This will further make the Sound Projector a 2D spatialisation instrument, easy to use for composers and artists.

Appendices

A Spherical coordinate system

The spherical coordinate system used throughout this work is defined as follows:

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = r \begin{pmatrix} \sin\vartheta \, \cos\varphi \\ \sin\vartheta \, \sin\varphi \\ \cos\vartheta \end{pmatrix} = r\,\boldsymbol{\theta}, \tag{A.1}$$

where the vector $\mathbf{r} \in \mathbb{R}^3$ denotes the radius, $\varphi \in [0, 2\pi)$ is the azimuth angle and $\vartheta \in [0, \pi]$ is the zenith angle. Note that the zenith angle ϑ is measured from the positive z coordinate. The inverse relation between spherical and Cartesian coordinates is given by

$$\begin{pmatrix} r \\ \varphi \\ \vartheta \end{pmatrix} = \begin{pmatrix} \sqrt{x^2 + y^2 + z^2} \\ \arctan(\frac{y}{x}) \\ \arctan(\frac{z}{r}) \end{pmatrix}. \tag{A.2}$$

The coordinate system is shown in fig. 39

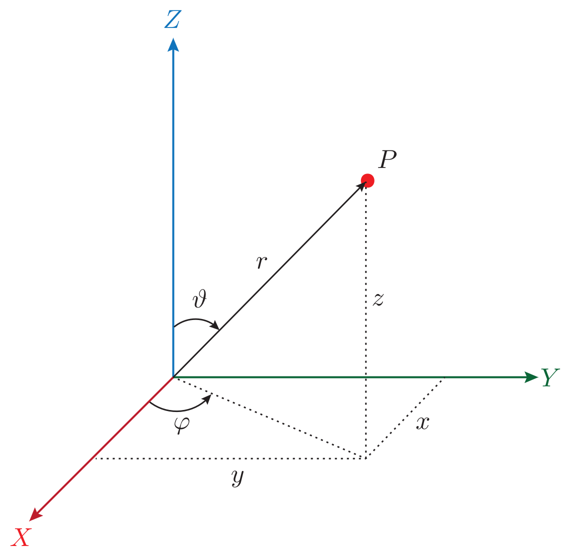


Figure 39: Definition of a point P in Cartesian and spherical coordinate system

B Loudspeaker Aperture Function

To compare the cap function $\boldsymbol{Y}_{cap}$ with the Dirac δ distribution $\boldsymbol{Y}_{\delta}$, a scaling factor N is derived from the Euclidean norm using the first 20 harmonics.

$$N = \frac{\|\boldsymbol{Y}_{cap}\|}{\|\boldsymbol{Y}_{\delta}\|}, \tag{B.1}$$

where $\|.\|$ denotes the Euclidean norm, Y_{Dirac} the SSH coefficients matrix of the 20 loudspeaker center positions and Y_{Cap} the SSH coefficients matrix of the cap function respectively. The relative error is calculated the following:

$$e = \frac{\boldsymbol{Y}_{cap} - N\,\boldsymbol{Y}_{\delta}}{\frac{\boldsymbol{Y}_{cap} + N\,\boldsymbol{Y}_{\delta}}{2}}\,100\% \tag{B.2}$$

Figure 40 and 41 show the error for each loudspeaker and SSH of order ν.

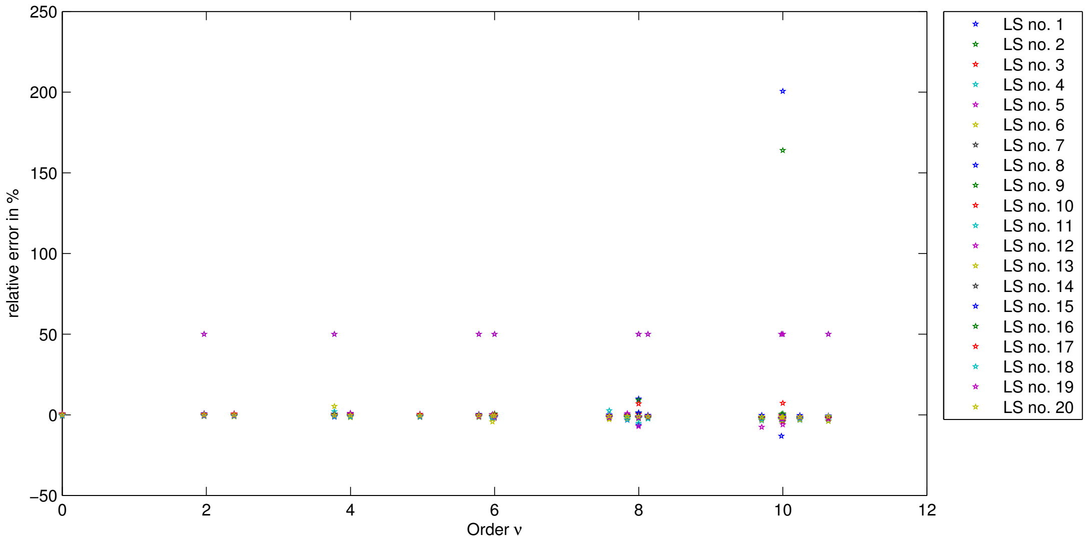


Figure 40: Error between cap functions $\boldsymbol{Y}_{cap}$ and Dirac δ distributions $\boldsymbol{Y}_{\delta}$

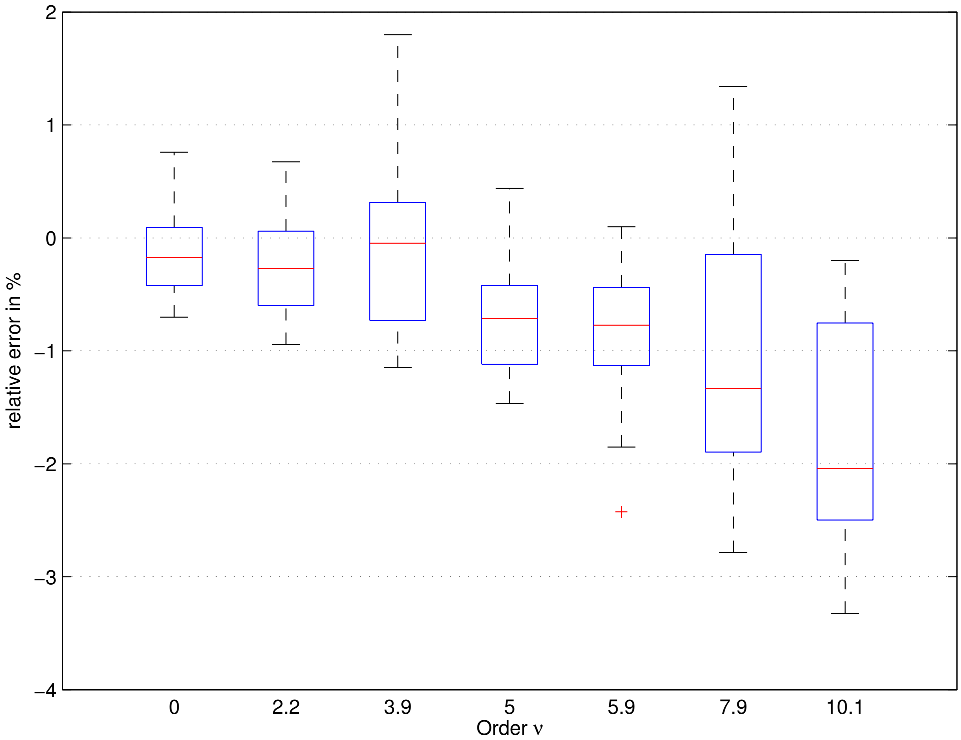


Figure 41: Error between cap functions $\boldsymbol{Y}_{cap}$ and Dirac δ distributions $\boldsymbol{Y}_{\delta}$ (outliers not considered), neighbouring orders ν are depicted grouped

References

[AB13] Haim Avron and Christos Boutsidis. Faster subset selections for matrices and applications. Technical report, Cornell University, 2013.

[AFKW06] Rimas Avizienis, Adrian Freed, Peter Kassakian, and David Wessel. A compact 120 independent element spherical loudspeaker array with programable radiation patterns. In *Audio Engineering Society Convention 120*, May 2006.

[Beh06] Gottfried K. Behler. Sound source with adjustable directivity. *The Journal of the Acoustical Society of America*, 120(5):3224–3224, 2006.

[DH94] J. R. Driscoll and D. M. Jr Healy. *Computing Fourier transforms and convolutions on the 2-sphere*, volume 15, pages 202–250. Adv. Appl. Math., 1994.

[Far00] Angelo Farina. Simultaneous measurement of impulse response and distortion with a swept-sine technique. In *108th AES Convention*, pages 18–22, 2000.

[GHB+05] Krzysztof M Gorski, Eric Hivon, AJ Banday, Benjamin D Wandelt, Frode K Hansen, Mstvos Reinecke, and Matthia Bartelmann. Healpix: a framework for high-resolution discretization and fast analysis of data distributed on the sphere. *The Astrophysical Journal*, 622(2):759, 2005.

[HC97] Cheinway Hwang and Shin-Kuen Chen. Fully normalized spherical cap harmonics: application to the analysis of sea-level data from topex/poseidon and ers-1. *Geophysical Journal International*, 129(2):450–460, 1997.

[Köß11] Thomas Kößler. Kugelkappenförmige Mikrofonanordnung zur 3-dimensionalen Erfassung des Schallfeldes. Master's thesis, IEM, 2011.

[KPPV14] Johannes Klein, Martin Pollow, Ander Perez, and Michael Vorländer. Measurement of room impulse responses with controllable directivity. *DAGA*, 2014.

[Kro14] Matthias Kronlachner. Spatial transformations for the alteration of ambisonic recordings. Master's thesis, IEM, 2014.

[KW04] Peter Kassakian and David Wessel. Characterization of spherical loudspeaker arrays. In *Audio Engineering Society Convention 117*, Oct 2004.

[Lös14] Stefan Lösler. MIMO-Rekursivfilter für Kugelarrays. Master's thesis, IEM, 2014.

[MBL07] Piotr Majdak, Peter Balazs, and Bernhard Laback. Multiple exponential sweep method for fast measurement of head-related transfer functions. *AES*, 2007.

[ME02] Jens Meyer and Gary Elko. A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield. In *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, volume 2, pages II–1781–II–1784, 2002.

[OS83] FWJ Olver and JM Smith. Associated legendre functions on the cut. *Journal of computational Physics*, 51(3):502–518, 1983.

[Pau13] Florian Pausch. A rigid double cone microphone array prototype. Master's thesis, IEM, 2013.

[PB09] Martin Pollow and Gottfried K. Behler. Variable directivity for platonic sound sources based on spherical harmonics optimization. *Acta Acustica united with Acustica*, 95(6):1082–1092, 2009.

[Pom10] Hannes Pomberger. Microphone arrays utilizing rigid cones for sound-field analysis. DAGA, 2010.

[PP14] Hannes Pomberger and Florian Pausch. Design and evaluation of a spherical segment array with double cone. *Proc. of the EAA Joint Symposium on Auralization and Ambisonics, Berlin, Germany, 3-5 April 2014*, 2014.

[PZ13] Hannes Pomberger and Franz Zotter. Modal sound field decomposition applicable for a limited range of directions. *Fortschritte der Akustik, AIA-DAGA, March 2013*, 2013.

[Raf05] B. Rafaely. Analysis and design of spherical microphone arrays. *Speech and Audio Processing, IEEE Transactions on*, 13(1):135–143, 2005.

[SZF14] Gerriet K Sharma, Franz Zotter, and Matthias Frank. Orchestrating wall reflections in space by icosahedral loudspeaker: findings from first artistic research exploration. *ICMC-SCM Athens*, 2014.

[WDC97] Olivier Warusfel, Philippe Derogis, and Rene Causse. Radiation synthesis with digitally controlled loudspeakers. In *Audio Engineering Society Convention 103*, Sep 1997.

[Wil99] Earl G. Williams. Fourier acoustics: Sound radiation and nearfield acoustical holography. 1999.

[ZFFR14] Franz Zotter, Matthias Frank, Andreas Fuchs, and Daniel Rudrich. Preliminary study on the perception of orientation-changing directional sound sources in rooms. In *Proc. of forum acusticum, Kraków*, 2014.

[ZH07] Franz Zotter and Robert Höldrich. Modeling radiation synthesis with spherical loudspeaker arrays. *ICA, Madrid*, 2007.

[ZN07] Franz Zotter and Markus Noisternig. Near- and farfield beamforming using spherical loudspeaker arrays. In *3rd Congress of the Alps Adria Acoustics Association*, 2007.

[Zot09] Franz Zotter. *Analysis and Synthesis of Sound-Radiation with Spherical Arrays*. PhD thesis, Kunstuniversität Graz, 2009.

[ZSH07] Franz Zotter, Alois Sontacchi, and Robert Höldrich. Modeling a spherical loudspeaker system as multipole source. *DAGA*, 2007.

[ZSN08] Franz Zotter, Andy Schmeder, and Markus Noisternig. Crosstalk cancellation for spherical loudspeaker arrays. *DAGA, Dresden*, 2008.